

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 25, 2004, 02:30:22 ; Search time 5273 Seconds  
(without alignments)  
10412.164 Million cell updates/sec

```
Title: US-10-621-787-1
Perfect score: 1161
Sequence: 1 ggaaactatgcctggggc.....aatccaacctttatcca 1161
```

```
scoring cable:  IDENT111_NUC
                  Gapop 10.0 , Gapext 1.0
```

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

# Search Notes

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1161	100.0	1161	6	AR399663	Sequence	
2	1161	100.0	1161	6	AX039655	Sequence	
3	1161	100.0	1161	6	AX698016	Sequence	
4	1161	100.0	1176	9	AF329836	Sequence	
5	1160	99.9	1157	9	AY528839	Sequence	
6	1154.2	99.4	1171	6	AX552224	Sequence	
7	1149	99.0	2395	6	BC011659	Sequence	
8	1137.2	98.0	1254	9	BC054506	Sequence	
9	783.8	67.5	152147	9	AC091842	Sequence	
c	10	783.8	67.5	153248	9	AC112191	Sequence
11	719.4	62.0	1236	10	BC030324	Mus musci	
12	702.2	60.5	1107	6	BD209657	Sequence	
13	702.2	60.5	1107	6	AR314173	Sequence	
14	701.6	60.4	1052	6	BD209734	Sequence	
15	701.6	60.4	1052	6	AR311540	Sequence	
16	627.4	54.0	886	6	BD211465	Sequence	
17	627.4	53.0	886	6	AX698018	Sequence	
18	615	54.0	972	6	CQ733580	Sequence	
19	613.6	52.9	855	6	AR339668	Sequence	

C	20	613.6	52.9	855	6	AX039974	Sequence
C	21	521.4	44.9	233609	10	AL670472	Mouse DNA
C	22	515.6	44.4	251117	2	AC0964835	Rattus no
C	23	507	43.7	855	6	BD211486	B211486 ACRP30 (3
C	24	473.8	40.8	2505	5	BX930999	Gallus ga
C	25	453	39.0	750	6	BD209547	Composblt1
C	26	453	39.0	750	6	AR341353	Sequence
C	27	338.4	29.1	409	11	G37397	SHGC-57747
C	28	337.4	29.1	409	6	AX054812	Sequence
C	29	309	26.6	1295	6	AX054815	Sequence
C	30	309	26.6	1295	10	BC038308	Mus muscu
C	31	295.4	25.4	912	6	AX054802	Sequence
C	32	295.4	25.4	1242	6	AX498268	Sequence
C	33	295.4	25.4	1306	9	AF329839	Homo sapi
C	34	295.4	25.4	3954	9	BC02187	Homo sapi
C	35	295.4	25.4	3954	9	BC024015	Homo sapi
C	36	292.2	25.2	4533	9	HSMB07927	Homo sapi
C	37	299.6	24.9	536	6	AR396689	Sequence
C	38	289.6	23.9	536	6	AX039975	Sequence
C	39	274.6	23.7	792	6	C0729382	Sequence
C	40	263.8	22.7	968	5	BX932235	Gallus ga
C	41	262.2	22.6	1662	5	EX950284	Gallus ga
C	42	261.6	22.5	225376	2	AC094230	Rattus no
C	43	261.6	22.5	233868	2	AC128545	Rattus no
C	44	254	21.9	187464	5	AL929007	Zebrafish
C	45	254	21.9	208135	5	EX004778	Zebrafish

## ALIGNMENTS

RESULT 1	AR399683	1161 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	AR399683				
DEFINITION	Sequence 1 from parent US 6620909.				
ACCESSION	AR399683				
VERSION	AR399683.1	GI:40142150			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 1161)				
AUTHORS	Piddington, C.S. and Bishop, P.D.				
TITLE	Adipocyte complement related protein homolog zacr2				
JOURNAL	Patent: US 6620909-A 1 16-SEP-2003;				
FEATURES	location/Qualifiers				
source	1..1161				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
Query Match	100.0%;	Score 1161;	DB 6;	Length 1161;	
Best Local Similarity	100.0%;	Pred. No. 3.1e-213;			
Matches 1161;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A	A
G	1	GGA	A	A	A
A	1	GGA	A	A	A
C	1	GGA	A	A	A
T	1	GGA	A	A</	

Db	241	AGCCTGCTGGCCCCCACAGGGCCACACCGGGCCCCCACAAGAGCCCAAGGGCCCTCAGAAATG	300
Qy	301	ATGGAGCGAATGGGCTTTCTTGGCAAGACGACCAAGATGCAACAGCGCGACCCGGGGG	360
Db	301	ATGGAGCGAATGGGCTTTCTTGGCAAGACGACCAAGATGCAACAGCGCGACCCGGGGG	360
Qy	361	GACAGCGGAGGAGAAAGGTCCACTGCGCGGACAGGTAAACCGGGGAAGCCAGACCAAG	420
Db	361	GACAGCGGAGGAGAAAGGTCCACTGCGCGGACAGGTAAACCGGGGAAGCCAGACCAAG	420
Qy	421	GGCAAAAGCCGGGGGCAATGGGCGGGCTTGGCCCCCGTGGCCCCCAAGGGGGTCAACGTTAC	480
Db	421	GGCAAAAGCCGGGGGCAATGGGCGGGCTTGGCCCCCGTGGCCCCCAAGGGGGTCAACGTTAC	480
Qy	481	CCCGGAAAGCATGGCAACAACGACGAAGAGGGCCCAAGGGCAAGAAAGGAGGACCAAGC	540
Db	481	CCCGGAAAGCATGGCAACAACGACGAAGAGGGCCCAAGGGCAAGAAAGGAGGACCAAGC	540
Qy	541	CTCCCAAGGCCCCCTGCACTGTGGCAATGGCCATACAAAGTCAAGCTTTCTCGGTGGCAATG	600
Db	541	CTCCCAAGGCCCCCTGCACTGTGGCAATGGCCATACAAAGTCAAGCTTTCTCGGTGGCAATG	600
Qy	601	ACCAAGAGCTACCCACGGGAGCGGGCTGGCCATCAAGTTTGAACAATTCATATGAACGAG	660
Db	601	ACCAAGAGCTACCCACGGGAGCGGGCTGGCCATCAAGTTTGAACAATTCATATGAACGAG	660
Qy	661	GGTGGCCACTCAAAATGCTTTCAGCGGGCAAGTTTGTCTGGCGGCGTGGCTGGGATCTTACAC	720
Db	661	GGTGGCCACTCAAAATGCTTTCAGCGGGCAAGTTTGTCTGGCGGCGTGGCTGGGATCTTACAC	720
Qy	721	TTTCACTTACGACATCAGCTGTGGCCACAAGCACTGTGGCATGGGCTGTGTGCAACAACGCGC	780
Db	721	TTTCACTTACGACATCAGCTGTGGCCACAAGCACTGTGGCATGGGCTGTGTGCAACAACGCGC	780
Qy	781	CAGTACCGCATCCGGGACCTTTGATGTGCCAAACCGGGCAACGACATGTGGGCTCCAGGCTCC	840
Db	781	CAGTACCGCATCCGGGACCTTTGATGTGCCAAACCGGGCAACGACATGTGGGCTCCAGGCTCC	840
Qy	841	ACCATCTGTGGCTTCAAGCAGGAGTGAACGAATTTGGCTGACAGATCTTCTACTCAGACGAG	900
Db	841	ACCATCTGTGGCTTCAAGCAGGAGTGAACGAATTTGGCTGACAGATCTTCTACTCAGACGAG	900
Qy	901	AACGGGCTTCTTATGACCTTTACTGGAACAACAGCTCTTTTACGGGCTTCTTAATCTAT	960
Db	901	AACGGGCTTCTTATGACCTTTACTGGAACAACAGCTCTTTTACGGGCTTCTTAATCTAT	960
Qy	961	GGCGACCAAGATGACCCCAACGAGGTATAGAATGSCACGGGGGCTCCACAGGACGGGAA	1020
Db	961	GGCGACCAAGATGACCCCAACGAGGTATAGAATGSCACGGGGGCTCCACAGGACGGGAA	1020
Qy	1021	CMACTTCTGACCTTGGGCTTACAGACCAAGCCCAACATGTAGGCTGGGGGATGGGGG	1080
Db	1021	CMACTTCTGACCTTGGGCTTACAGACCAAGCCCAACATGTAGGCTGGGGGATGGGGG	1080
Qy	1081	GTCAGATGAGCGGTTTCTAGCCTCAGGGCTCACCTCTCGGCTCTTTTTCCTCCCTTCAAT	1140
Db	1081	GTCAGATGAGCGGTTTCTAGCCTCAGGGCTCACCTCTCGGCTCTTTTTCCTCCCTTCAAT	1140
Qy	1141	AAATCCAAACCTTTTATTTCA 1161	
Db	1141	AAATCCAAACCTTTTATTTCA 1161	

RESULT 2			
LOCUS	AX039965		
DEFINITION	AX039965	1161 bp	DNA
ACCESSION	Sequence 1 from Patent WO006376.		linear
VERSION	AX039965		
KEYWORDS	AX039965.1 GI:1122984		
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Piddington,C.S. and Bishop,P.D.
TITLE	Adipocyte complement related protein homolog zacrp2
JOURNAL	Patent: WO 0063376-A 1 26-OCT-2000; ZymoGenetics, Inc. (us)
FEATURES	location/qualifiers 1..1161
SOURCE	/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 133..990 /note="unnamed protein product" /codon_start=1 /protein_id="CACI6634.1" /db_xref="GI:11229985" /translation="MIPIPVLLACALPCADPLLGAFARDPFRKSPOLVCSLIPGPGK PGRPARPGSGMMGRMGPRKDQDGDGDEGGSEBGPRTGNRGRPKRGXAGA GRAPRPARKYNTPRGHGTGPRGSKGPRDKGEPELPGPSGSQHTTSAFSLVALTKSY PREELPTIKFDILINBEIGHIVASSGKVCGVGLYYFTIDTLANKLATIGLVHQO RIIRFPDANTGMHDVASGSTLIALKQGBEWMQLIFYSEONGLFDPYPWDSLTGFLIT ADQDDPNBV"
CDS	
ORIGIN	
Query Match	100.0%; Score 1161; DB 6; Length 1161;
Best Local Similarity	100.0%; Prid. No. 3.1e-213;
Matches 1161; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GGAAATCATAGCCCTGGGGCCGACGCCTGCCCCGGCTGCGCCTGAGAAAAGCCGGAC 60
DB	1 GGAAAACTAATCTCTGGGGCCCAACGCTTGCCCCGGCTGCTGCCCTGAGAAAAGCCGGAAC 60
OY	61 GCGAGGCCCGCCGAGAGCTTTCTTGTCTCGGAACGCCCTGTGACGTVGGCGGGACGCGG 120
DB	61 GCGAGGCCCGCCGAGAGCTTTCTTGTCTCGGAACGCCCTGTGACGTVGGCGGGACGCGG 120
OY	121 AGGCTAACCAACAATGATTCCTCGTGGTGTCTCTGGCTGTGCTTCCCTGTGCTGTAC 180
DB	121 AGGGTAACCAACAATGATTCCTCGTGGTGTCTCTGGCTGTGCTTCCCTGTGCTGTAC 180
OY	181 CCACGTGCTTGGCGGCTTTGTCTCGAGGGGACTTCGCGAAAAGGCTCCCTCAACTGGTCTGC 240
DB	181 CCACTGTCTTGGCGGCTTTGTCTCGAGGGGACTTCGCGAAAAGGCTCCCTCAACTGGTCTGC 240
OY	241 AGCCTGCTTGGCCCCCAGAGGCCCAACCCGGCCCCCAGAGAGCCCAGAGGCTTCAAGAAATG 300
DB	241 AGCCTGCTTGGCCCCCAGAGGCCCAACCCGGCCCCCAGAGAGCCCAGAGGCTTCAAGAAATG 300
OY	301 ATGGAGCAATATGGGCTTCTCTGGCAAAGACGGCCAATGACACGACGCGGACCGGGGG 360
DB	301 ATGGGACCAATATGGGCTTCTCTGGCAAAGACGGCCAATGACACGACGCGGACCGGGGG 360
OY	361 GACAGCGGAGAGAAAGTCCACTGGCCCGGACAGGTAAACGGGGGAAAGCCAGACCAAG 420
DB	361 GACAGCGGAGAGAAAGTCCACTGGCCCGGACAGGTAAACGGGGGAAAGCCAGACCAAG 420
OY	421 GGCAAAAGCCGGGGCAATTGGGCGGGCTTGGCCCCCGTGGCCCCCAAGGGGGTCAACGTACC 480
DB	421 GGCAAAAGCCGGGGCAATTGGGCGGGCTTGGCCCCCGTGGCCCCCAAGGGGGTCAACGTACC 480
OY	481 CCCGGGAAGCATGGCAACAACAGGAGAGGGGCCCAAGGGCCAGAAAGGGGAGCCAGGC 540
DB	481 CCCGGGAAGCATGGCAACAACAGGAGAGGGGCCCAAGGGCCAGAAAGGGGAGCCAGGC 540
OY	541 CTCCCAAGGCCCCCTGACGTGTGGGCAATGCGCATACAAATCAAGTCAAGTTTCTCGTGGCAATG 600
DB	541 CTCCCAAGGCCCCCTGACGTGTGGGCAATGCGCATACAAATCAAGTCAAGTTTCTCGTGGCAATG 600
OY	601 ACCAAGAGCTACCAACGGGAGCGGCTGGCCATCAAGTTTGAACAAGATTCTGATGAACAG 660
DB	601 ACCAAGAGCTACCAACGGGAGCGGCTGGCCATCAAGTTTGAACAAGATTCTGATGAACAG 660
OY	661 GGTGGCCACTACAAATGCTTCCAGAGGCAAGTTTGGTGTGGCGGCTGGGATCTTAATAC 720

Db 661 GGTGGCCATCTACATGCTTCAGAGCGGCAAGTTCGTCTGCGGCTGCTGGATCTACTAC 720  
Qy 721 TTCACTACGACATACGCTGGCCACAGACCTGGCCATCGGCTGGTGCACACGGC 780  
Db 721 TTCACTACGACATACGCTGGCCACAGACCTGGCCATCGGCTGGTGCACACGGC 780  
Qy 781 CAGTACCGCATCCGGACCTTTGATGACCAACCGGCAACGATGGGCTCAGGCTCC 840  
Db 781 CAGTACCGCATCCGGACCTTTGATGACCAACCGGCAACGATGGGCTCAGGCTCC 840  
Qy 841 ACCATCTGCTCTCAAGCAGAGGTGACGAGTTGGCTGACATCTTCTACTCAGACAG 900  
Db 841 ACCATCTGCTCTCAAGCAGAGGTGACGAGTTGGCTGACATCTTCTACTCAGACAG 900  
Qy 901 AACGGGCTCTTCTATGACCTTACTGACAGACGCTCTTTACGGGCTCTTATCTAT 960  
Db 901 AACGGGCTCTTCTATGACCTTACTGACAGACGCTCTTTACGGGCTCTTATCTAT 960  
Qy 961 GCCGACGAGATGACCCCAACGAGTATGACATGCCAGGCGGCTCTCCAGAGGAGAA 1020  
Db 961 GCCGACGAGATGACCCCAACGAGTATGACATGCCAGGCGGCTCTCCAGAGGAGAA 1020  
Qy 1021 CAAGCTCTGGACTTGGGCTTACAGAGCAAGACCCCAACTGTAGGCTGGGCGTGGGCG 1080  
Db 1021 CAAGCTCTGGACTTGGGCTTACAGAGCAAGACCCCAACTGTAGGCTGGGCGTGGGCG 1080  
Qy 1081 GTGAGTGAAGGGTCTTAGAGCTCAAGCTCCTCGGCTCTTTTTCCTCCCTCAT 1140  
Db 1081 GTGAGTGAAGGGTCTTAGAGCTCAAGCTCCTCGGCTCTTTTTCCTCCCTCAT 1140  
Qy 1141 AATCCAACTTTTATCA 1161  
Db 1141 AATCCAACTTTTATCA 1161

RESULT 3  
AX698016  
LOCUS AX698016 1161 bp DNA linear PAT 02-APR-2003  
DEFINITION Sequence 1 from Patent WO03010197.  
ACCESSION AX698016  
VERSION AX698016.1 GI:29499054  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Lucas, J., Scallia, A. and Dialynas, D.  
TITLE Gng-1 polynucleotides and polypeptides and uses thereof  
JOURNAL Patent: WO 03010197-A 1 06-FEB-2003;  
GENSET (FR)  
FEATURES  
source 1. 1161  
Location/Qualifiers  
1. 1161  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1. 1132  
133..990  
/note="unamed protein product"  
/codon\_start=1  
/protein\_id="CAD87410.1"  
/db\_xref="GI:29499054"  
/translation="MIPWVLLCALPCADPELGAFAARDFRKSPOLVCSLPQPGP  
PGPGRAPGSGMGRMGFPGKDGSDGHDGDSGSGSPRTGNRKGPGKAGAI  
GRAPGRPGKVGTPGKGTGPKGKGPGLPBGCSGSHTSASFVAVTKSY  
PRRLPIKPKILMNEGHYVASGKPVCGVPGIYFTYDITLANKILATGLVNGY  
RIKRPDANTGHDVASSSTILAKQGBVMILQIRYSRNGLFPIYPTDLSFTGFLY  
ADDDPENSV"  
CDS  
5' UTR  
3' UTR  
ORIGIN  
3' UTR  
991..1161

Query Match 100.0%; Score 1161; DB 6; Length 1161;

Best Local Similarity 100.0%; Pred. No. 3, 1e-213;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAATCTATGCTGGGCGGAGGCTTGCCCGGCTGCTGCGTGAAGAAAGCCGGAGC 60  
Db 1 GAAATCTATGCTGGGCGGAGGCTTGCCCGGCTGCTGCGTGAAGAAAGCCGGAGC 60  
Qy 61 GCGAGGCGGCGGAGGCTTCTTGGCTCGGAGCGGCCCTGGAGCGTGGCGGAGCGCGG 120  
Db 61 GCGAGGCGGCGGAGGCTTCTTGGCTCGGAGCGGCCCTGGAGCGTGGCGGAGCGCGG 120  
Qy 121 AGGGTAAACAACATGATCCCTGGGTGCTCTGAGCTGTGACCTCCCTGTGCTGAC 180  
Db 121 AGGGTAAACAACATGATCCCTGGGTGCTCTGAGCTGTGACCTCCCTGTGCTGAC 180  
Qy 181 CCACTGCTTGGCGCTTGTGCTCGAGGGAATTCCGAAAGGCTCCCTCAATGCTGTC 240  
Db 181 CCACTGCTTGGCGCTTGTGCTCGAGGGAATTCCGAAAGGCTCCCTCAATGCTGTC 240  
Qy 241 AGCTGCTGGCCCCGAGGGCCACCGGGCCCCAGAGGCCCGAGGCGCTCAGAAATG 300  
Db 241 AGCTGCTGGCCCCGAGGGCCACCGGGCCCCAGAGGCCCGAGGCGCTCAGAAATG 300  
Qy 301 ATGGAGCAATGAGCTTCTTGGCAAGAGCGCCAGATGAGACAGAGCGGAGC 360  
Db 301 ATGGAGCAATGAGCTTCTTGGCAAGAGCGCCAGATGAGACAGAGCGGAGC 360  
Qy 361 GACAGCGAGAGAGAGGTCCACTGCGCGAGCAGGTAAACGGGGAAAGCCAGGAC 420  
Db 361 GACAGCGAGAGAGAGGTCCACTGCGCGAGCAGGTAAACGGGGAAAGCCAGGAC 420  
Qy 421 GGCAGAGCGGGGCAATTGGCGGGGCTGGGCCCGTGGCCCCCAAGGGGGTCAACGGTAC 480  
Db 421 GGCAGAGCGGGGCAATTGGCGGGGCTGGGCCCGTGGCCCCCAAGGGGGTCAACGGTAC 480  
Qy 481 CCCGGAGATGGCACAACGAGCAAGAGGAGCCCAAGGAGCAAGAAAGGAGACCAAGC 540  
Db 481 CCCGGAGATGGCACAACGAGCAAGAGGAGCCCAAGGAGCAAGAAAGGAGACCAAGC 540  
Qy 541 CTCCAGAGCCCTGACGCTGTGGCAATGCGCATACAAATGAGCTTCTGCTGGGCGATG 600  
Db 541 CTCCAGAGCCCTGACGCTGTGGCAATGCGCATACAAATGAGCTTCTGCTGGGCGATG 600  
Qy 601 ACCAAGAGCTACCAACGAGGAGCGGCTGCCATGAATTCAGAAATTCGATGAAGAG 660  
Db 601 ACCAAGAGCTACCAACGAGGAGCGGCTGCCATGAATTCAGAAATTCGATGAAGAG 660  
Qy 661 GGTGGCACTACATGCTTCCAGGCGCAAGTTGCTGCGGCGTGGCGATCTACTAC 720  
Db 661 GGTGGCACTACATGCTTCCAGGCGCAAGTTGCTGCGGCGTGGCGATCTACTAC 720  
Qy 721 TTCACTACGACATACGCTGGCCACAGACCTGGCCATCGGCTGGTGCACACGGC 780  
Db 721 TTCACTACGACATACGCTGGCCACAGACCTGGCCATCGGCTGGTGCACACGGC 780  
Qy 781 CAGTACCGCATCCGGACCTTTGATGACCAACCGGCAACGATGGGCTCAGGCTCC 840  
Db 781 CAGTACCGCATCCGGACCTTTGATGACCAACCGGCAACGATGGGCTCAGGCTCC 840  
Qy 841 ACCATCTGCTCTCAAGCAGAGGTGACGAGTTGGCTGACATCTTCTACTCAGACAG 900  
Db 841 ACCATCTGCTCTCAAGCAGAGGTGACGAGTTGGCTGACATCTTCTACTCAGACAG 900  
Qy 901 AACGGGCTCTTCTATGACCTTACTGACAGACGCTCTTTACGGGCTCTTATCTAT 960  
Db 901 AACGGGCTCTTCTATGACCTTACTGACAGACGCTCTTTACGGGCTCTTATCTAT 960  
Qy 961 GCCGACGAGATGACCCCAACGAGTATGACATGCCAGGCGGCTCTCCAGAGGAGAA 1020  
Db 961 GCCGACGAGATGACCCCAACGAGTATGACATGCCAGGCGGCTCTCCAGAGGAGAA 1020  
Qy 1021 CAAGCTCTGGACTTGGGCTTACAGAGCAAGACCCCAACTGTAGGCTGGGCGTGGGCG 1080

Db 1021 CAAGCTTCTGGACTTGGGCTTACAGAGCAAGACCACAACCTAGGCTGGGGGTGGGG 1080  
Qy 1081 GTGAGAGGAGCGGTTTACAGCTGAGGCTGACCTCTCGCTCTTTTTCCTCCCTTCAAT 1140  
Db 1081 GTGAGAGGAGCGGTTTACAGCTGAGGCTGACCTCTCGCTCTTTTTCCTCCCTTCAAT 1140  
Qy 1141 AAATCCAAACCTTTTATTTCA 1161  
Db 1141 AAATCCAAACCTTTTATTTCA 1161

RESULT 4  
AF329836 1176 bp mRNA linear PRI 12-MAR-2001  
LOCUS Homo sapiens complement-clq tumor necrosis factor-related protein  
DEFINITION (CTRP2) mRNA, complete cds.  
ACCESSION AF329836  
VERSION AF329836.1 GI:13274517  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1176)  
AUTHORS Piddington,C.S. and Bishop,P.  
TITLE Homo sapiens complement-clq tumor necrosis factor-related protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1176)  
AUTHORS Piddington,C.S. and Bishop,P.  
TITLE Direct Submission  
JOURNAL Submitted (20-DEC-2000) Bioinformatics, ZymoGenetics, Inc., 1201  
Bastlake Ave. East, Seattle, WA 98102, USA

FEATURES  
source location/Qualifiers  
1..1176  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5q34"  
1..1176  
/gene="CTRP2"  
/note="ZACRP2"  
148..1005  
/gene="CTRP2"  
/codon\_start=1  
/product="complement-clq tumor necrosis factor-related  
protein"  
/protein\_id="AAK17960.1"  
/db\_xref="GI:13274518"  
/translation="MIPVLLACALPCADPLGAFARBDPRKGSPLVCSLPQPS  
RGPAGAPSGMMRMGFPKQDQDHDGRGDSBERGPRGRTGNRKPQKAGAI  
GRAGPRPKVNGTRPKRGTPKRGKGRKGRBRLPRCCSGSHRTSAISVATKSY  
PRSLPTKFDILINBEGHTNASSGKVCVPGIYTFYITLLANKRLAIGLVHNGY  
RIKFDANTGNHDVAGSTIIALKQGBEVMQLFYSEQNLFFDPYWTDSLFTGFLY  
ADDDPDEMEV"

ORIGIN  
Query Match 100.0%; Score 1161; DB 9; Length 1176;  
Best Local Similarity 100.0%; Pred. No. 3.1e-213;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAAAATATATGCTGGGGGCGAGAGCTGTGCGGCTGTGCGGCTGAGAGAAAGCGGGAGC 60  
Db 16 GGAATACTATGCTGGGGGCGAGAGCTGTGCGGCTGTGCGGCTGAGAGAAAGCGGGAGC 75  
Qy 61 GCGGAGCCCGCGGAGAGCTTTGCTCCGAGAGCCCTTGAAGTGGCGGGCAGCCGCG 120  
Db 76 GCGGAGCCCGCGGAGAGCTTTGCTCCGAGAGCCCTTGAAGTGGCGGGCAGCCGCG 135  
Qy 121 AGGCTAACCAACATGATCCCTGGGTGCTCTGCGCTGTGCTCCCTGTGCTGTGAC 180  
Db 136 AGGCTAACCAACATGATCCCTGGGTGCTCTGCGCTGTGCTCCCTGTGCTGTGAC 195

Qy 181 CCACTGCTTGGGCGCTTTTGCTCCGAGGAACTTCCGGAAGAGCTCCCTCAACTGTGTGC 240  
Db 196 CCACTGCTTGGGCGCTTTTGCTCCGAGGAACTTCCGGAAGAGCTCCCTCAACTGTGTGC 255  
Qy 241 AGCTGCTTGGGCGCTTTTGCTCCGAGGAACTTCCGGAAGAGCTCCCTCAACTGTGTGC 300  
Db 256 AGCTGCTTGGGCGCTTTTGCTCCGAGGAACTTCCGGAAGAGCTCCCTCAACTGTGTGC 315  
Qy 301 ATGGAGCGAATGAGGCTTTTCTGCAAGAGAGCGCAAGATGGAACAAGCGGCGA 360  
Db 316 ATGGAGCGAATGAGGCTTTTCTGCAAGAGAGCGCAAGATGGAACAAGCGGCGA 375  
Qy 361 GACAGCGAAGAGAGAGTCCACTGCGGAGCAGAGTAAACCGGGGAAACCGAGACCAAG 420  
Db 376 GACAGCGAAGAGAGAGTCCACTGCGGAGCAGAGTAAACCGGGGAAACCGAGACCAAG 435  
Qy 421 GGCAGAGCCGGGGCCATTGGGCGGGCTGGCCCCCGTGGCCCAAGGGGGTCAACGGTACC 480  
Db 436 GGCAGAGCCGGGGCCATTGGGCGGGCTGGCCCCCGTGGCCCAAGGGGGTCAACGGTACC 495  
Qy 481 CCGGGAGAGATGGGCAACAGGCGAAGAGGGGCCAAGGGGCAAGAGGGGAGCGAGC 540  
Db 496 CCGGGAGAGATGGGCAACAGGCGAAGAGGGGCCAAGGGGCAAGAGGGGAGCGAGC 555  
Qy 541 CTCCAGAGCCCTGACAGTGTGGCAGTGGCCATACCAAGTCAAGCTTTCTCGTGGCAGTG 600  
Db 556 CTCCAGAGCCCTGACAGTGTGGCAGTGGCCATACCAAGTCAAGCTTTCTCGTGGCAGTG 615  
Qy 601 ACCAAGAGTACCAACGAGGAGCGGGCTGGCCCATCAAGTTTGAACAAGTTCTGATGAACGAG 660  
Db 616 ACCAAGAGTACCAACGAGGAGCGGGCTGGCCCATCAAGTTTGAACAAGTTCTGATGAACGAG 675  
Qy 661 GGTGGCCATCAATGCTTCCAGGGCAAGTTGCTGGGGCGGTGCTGGAGTCTACTAC 720  
Db 676 GGTGGCCATCAATGCTTCCAGGGCAAGTTGCTGGGGCGGTGCTGGAGTCTACTAC 735  
Qy 721 TTCACTTACGACATCAAGCTGTGGCCCAAGAGCACTGGCCCTGTGTGCAACGAGC 780  
Db 736 TTCACTTACGACATCAAGCTGTGGCCCAAGAGCACTGGCCCTGTGTGCAACGAGC 795  
Qy 781 CAGTACCGCATCCGAGACCTTTGATGCCAACACCGGCAACAGATGTGGCTCAAGCTCC 840  
Db 796 CAGTACCGCATCCGAGACCTTTGATGCCAACACCGGCAACAGATGTGGCTCAAGCTCC 855  
Qy 841 ACCATCTGGCTGCAAGAGGAGGTGAAGAGTTGGCTGAGATCTTCACTCAGAGCAG 900  
Db 856 ACCATCTGGCTGCAAGAGGAGGTGAAGAGTTGGCTGAGATCTTCACTCAGAGCAG 915  
Qy 901 AACGGCTCTTCTATGACCTTTACTGAGACAGAGCTCTTTACGGGCTTCTAATCTAT 960  
Db 916 AACGGCTCTTCTATGACCTTTACTGAGACAGAGCTCTTTACGGGCTTCTAATCTAT 975  
Qy 961 GCCGACCAAGATGACCCCAAGAGATGACATGCCACGGCGGTCTCTCAGGCAAGGAA 1020  
Db 976 GCCGACCAAGATGACCCCAAGAGATGACATGCCACGGCGGTCTCTCAGGCAAGGAA 1035  
Qy 1021 CAAGCTTCTGGGACTTGGGCTTACAGAGCAAGACCCCAAACTGAGGCTGGGGTGGGG 1080  
Db 1036 CAAGCTTCTGGGACTTGGGCTTACAGAGCAAGACCCCAAACTGAGGCTGGGGTGGGG 1095  
Qy 1081 GTGAGTGAAGCGGTTTACAGCTGAGGCTCACTCTCGGCTCTTTTTCCTCCCTCAAT 1140  
Db 1096 GTGAGTGAAGCGGTTTACAGCTGAGGCTCACTCTCGGCTCTTTTTCCTCCCTCAAT 1155  
Qy 1141 AAATCCAAACCTTTTATTTCA 1161  
Db 1156 AAATCCAAACCTTTTATTTCA 1176

RESULT 5  
AY358839 1167 bp mRNA linear PRI 03-OCT-2003  
LOCUS Homo sapiens clone DNA143501 CTRP2 (UNQ6349) mRNA, complete cds.  
DEFINITION



ACCESSION AY358839.1 GI:37182795  
VERSION AY358839.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1167)  
AUTHORS Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J., Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B., Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Haas,P.E., Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robble,E., Sanchez,C., Schoenfeld,J., Seshagiri,S., Simone,L., Singh,J., Smith,V., Stinson,J., Vagstad,A., Vandlen,R., Watanabe,C., Weiland,D., Woods,K., Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,W.I. and Godowski,P.  
The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: A Bioinformatics Assessment  
Genome Res. 13 (10), 2265-2270 (2003)  
JOURNAL  
PUBMED 12975309  
REFERENCE 2 (bases 1 to 1167)  
AUTHORS Clark,H.F.  
TITLE Direct Submission  
JOURNAL Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA  
FEATURES  
source  
1..1167  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DNA143501"  
1..1167  
/locus\_tag="UNQ6349"  
140..597  
/locus\_tag="UNQ6349"  
/notes="PRO21054"  
/product="CTRP2"  
/protein\_id="AA089198.1"  
/db\_xref="GI:37182796"  
/translation="MIPWVLLALPCADPLIGAPARDRKSPOLVCSLSPGPGP  
PGPRGAPRSGMGRMRKPGKDGDDHGDHGDGDSGSPRRTNRKRGPKKAKAI  
GRAGPRKGVNGTPKRGKPGKGPGLPGPCSGSGSHTSASVAVTXSY  
PRRLPIKFDKILNKGHNVAASGKGVCGPIYFTYDITLANKHIAIGLVNIGY  
RIRTFDANTGNHVDVAGSTIILAKQDSEVWLQIFSRQNGLFYDPYTDLSLFTFLY  
ADQDDPNEV"  
ORIGIN  
Query Match 99.9%; Score 1160; DB 9; Length 1167;  
Best Local Similarity 100.0%; Pred. No. 4.8e-213;  
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGAATATGAGCTGGGGCGGAGCGCTGCGCGCGCTGCGCGGAGAAAGCGCGAGC 60  
DB 8 GGAATATGAGCTGGGGCGGAGCGCTGCGCGCGCTGCGCGGAGAAAGCGCGAGC 67  
QY 61 GCGAGCGCGCGCGAGAGCTTCTTGTCTCGGAGCGCGCGCTGAGAGTGGCGGAGCGCG 120  
DB 68 GCGAGCGCGCGCGAGAGCTTCTTGTCTCGGAGCGCGCGCTGAGAGTGGCGGAGCGCG 127  
QY 121 AGGGTAACACATGATCCCTGGGATGCTCTGTGCTGTGCTTCCCTGTGCTGAC 180  
DB 128 AGGGTAACACATGATCCCTGGGATGCTCTGTGCTGTGCTTCCCTGTGCTGAC 187  
QY 181 CCACGTGCTGGCGCTTGTCTGCGAGGGAGCTTCCGGAAGGCTCCCTCAACTGTCTGC 240  
DB 188 CCACGTGCTGGCGCTTGTCTGCGAGGGAGCTTCCGGAAGGCTCCCTCAACTGTCTGC 247  
QY 241 AGCTGCTGGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
DB 248 AGCTGCTGGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 307

QY 301 ATGGAGCAATGGGCTTCTCTGGCAAGAGCGGCCAAGATGGACACAGCGCGACCGGGGG 360  
DB 308 ATGGAGCAATGGGCTTCTCTGGCAAGAGCGGCCAAGATGGACACAGCGCGACCGGGGG 367  
QY 361 GACAGCGAGAGAGAGAGGTTCACCTGCGCGGACAGGTAAACGGGGAAACCAAGCAAG 420  
DB 368 GACAGCGAGAGAGAGAGGTTCACCTGCGCGGACAGGTAAACGGGGAAACCAAGCAAG 427  
QY 421 GCGAAGCGCGGGCATTTGGCGGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
DB 428 GCGAAGCGCGGGCATTTGGCGGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 487  
QY 481 CCGGAGAGAGATGGACACACAGCGCAAGAGGGGCCAAGGGGCAAGAGGGGCGCGAGC 540  
DB 488 CCGGAGAGAGATGGACACACAGCGCAAGAGGGGCCAAGGGGCAAGAGGGGCGCGAGC 547  
QY 541 CTCCAGAGCGCGCGAGCTGCGAGTGGGACATACCAAGTCACTTCTCGTGGCAGTG 600  
DB 548 CTCCAGAGCGCGCGAGCTGCGAGTGGGACATACCAAGTCACTTCTCGTGGCAGTG 607  
QY 601 ACCAAGAGCTAACCCACGAGCGAGCGGCTGCCATCAAGTTTGAACAAGATTGATGACGAG 660  
DB 608 ACCAAGAGCTAACCCACGAGCGAGCGGCTGCCATCAAGTTTGAACAAGATTGATGACGAG 667  
QY 661 GGTGGCCATCAATGCTTTCAGCGGCAAGTTGTCGCGCGCGTGGCGGATCTACTAC 720  
DB 668 GGTGGCCATCAATGCTTTCAGCGGCAAGTTGTCGCGCGCGTGGCGGATCTACTAC 727  
QY 721 TTCACTAGACATCACTGCGCGCAACCAAGCACTGCGCGCTGCGTGCACACGCG 780  
DB 728 TTCACTAGACATCACTGCGCGCAACCAAGCACTGCGCGCTGCGTGCACACGCG 787  
QY 781 CAGTACCGCATTCGAGACCTTTGATGCAACACCGGCAACCAAGTGGCTCAAGCTCC 840  
DB 788 CAGTACCGCATTCGAGACCTTTGATGCAACACCGGCAACCAAGTGGCTCAAGCTCC 847  
QY 841 ACCATCTGCGCTCTCAAGAGAGGAGAGCAAGTTGCTGAGATCTTCACTACAGAGCAG 900  
DB 848 ACCATCTGCGCTCTCAAGAGAGGAGAGCAAGTTGCTGAGATCTTCACTACAGAGCAG 907  
QY 901 AACGGGCTTCTATGACCTCTTACTGACAGAGCTCTTCAAGGCTTCTATATCTAT 960  
DB 908 AACGGGCTTCTATGACCTCTTACTGACAGAGCTCTTCAAGGCTTCTATATCTAT 967  
QY 961 GCGACCAAGATGACCCCAAGAGTATGACATGCAACGCGCGCTCTCAAGGAGGAA 1020  
DB 968 GCGACCAAGATGACCCCAAGAGTATGACATGCAACGCGCGCTCTCAAGGAGGAA 1027  
QY 1021 CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGACCCCAACTGATGGCTGGGGTGGGG 1080  
DB 1028 CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGACCCCAACTGATGGCTGGGGTGGGG 1087  
QY 1081 GTGAGTGAAGCGGTTCTAGGCTCAAGGCTCACTCTCGGCTCTTTTTCCTTCAAT 1140  
DB 1088 GTGAGTGAAGCGGTTCTAGGCTCAAGGCTCACTCTCGGCTCTTTTTCCTTCAAT 1147  
QY 1141 AATTCCAACCTTTTATTC 1160  
DB 1148 AATTCCAACCTTTTATTC 1167  
RESULT 6  
AX552234  
LOCUS AX552234 1171 bp DNA linear PAT 27-NOV-2002  
DEFINITION Sequence 143 from Patent WO0162927.  
ACCESSION AX552234  
VERSION AX552234.1 GI:25896472  
KEYWORDS  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
1  
Banville, S.C., Greenawalt, L.B., Lincoln, S.E., Stockdreher, T.K.,  
Anshev, S., Chang, S.C., Chen, W., D'Sa, S.A., Dam, T.C., Liu, T.F.,  
Rosen, B.H., Russo, F.D., Spiro, P.A., Bradley, D.L., Chen, A.,  
Chen, H.T., Datto, A., Daniels, S.B., Dutoit, G.B., Flores, V.,  
Fong, W.T., Hodgson, D.M., Jackson, S., Jones, A.L., Panzer, V.,  
Roseberry, A.M., Shah, P., Wright, R.J., Yap, P.B., Yu, J.Y.,  
Bratcher, S.R., Chalup, M.S., Dahl, C.R., and Hillman, J.L.  
Polypeptides and corresponding polynucleotides for diagnostics and  
therapeutics  
Patent: WO 0162927-A 143 30-AUG-2001;  
JOURNAL  
Incyte Genomics, Inc. (US)  
Location/Qualifiers  
FEATURES  
source  
1. 1171  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: LI:414307.1:2000FEB01"

ORIGIN  
Query Match 99.4%; Score 1154.2; DB 6; Length 1171;  
Best Local Similarity 99.6%; Pred. No. 6.3e-212;  
Matches 1156; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GAAAATATATGCTGGGCGCGAGCTCTGCGCGCTGCTGAGAGAAAGCGGAGAC 60  
DB 8 GAAAATATATGCTGGGCGCGAGCTCTGCGCGCTGCTGAGAGAAAGCGGAGAC 67  
QY 61 GCGGAGCGCGCGCGAGAGCTTTTGTCTCGGAGAGCGCGCTGAGAGTGGCGGCGCG 120  
DB 68 GCGGAGCGCGCGCGAGAGCTTTTGTCTCGGAGAGCGCGCTGAGAGTGGCGGCGCG 127  
QY 121 AGGGTAACACCATGATCCCTGGGATGCTCTGCGCTGCGCGCTGCGCTGAGTGC 180  
DB 128 AGGGTAACACCATGATCCCTGGGATGCTCTGCGCTGCGCGCTGCGCTGAGTGC 187  
QY 181 CCATGCTTGGCGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAATGTCG 240  
DB 188 CCATGCTTGGCGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAATGTCG 247  
QY 241 AGCTGCTTGGCGCGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAATGTCG 300  
DB 248 AGCTGCTTGGCGCGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAATGTCG 307  
QY 301 ATGGGAGCAATGCGCTTGTCTGCGAAAGAGCGGCAAGTGAACAGCGGCGG 360  
DB 308 ATGGGAGCAATGCGCTTGTCTGCGAAAGAGCGGCAAGTGAACAGCGGCGG 367  
QY 361 GACAGCGGAGAGAGAGGTCCACTGCGCGGAGAGTAAACCGGGGAAAGCCAGCAAG 420  
DB 368 GACAGCGGAGAGAGAGGTCCACTGCGCGGAGAGTAAACCGGGGAAAGCCAGCAAG 427  
QY 421 GCGAAGACCGGGGCGCATTTGGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCG 480  
DB 428 GCGAAGACCGGGGCGCATTTGGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCG 487  
QY 481 CCCGGAGAGATGCGACACAGCAAGAGAGGCGCCCAAGGCGCAAGAGAGGCGGCG 540  
DB 488 CCCGGAGAGATGCGACACAGCAAGAGAGGCGCCCAAGGCGCAAGAGAGGCGGCG 547  
QY 541 CTCCTCAGCGCGCTGCGAGCTGCGAGTGGCCATCAAGTCAAGTTCCTCGTGGCAGTG 600  
DB 548 CTCCTCAGCGCGCTGCGAGCTGCGAGTGGCCATCAAGTCAAGTTCCTCGTGGCAGTG 607  
QY 601 ACCAAGAGCTTACCAAGGAGAGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 660  
DB 608 ACCAAGAGCTTACCAAGGAGAGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 667  
QY 661 GGTGGCACTACATGCTTCCAGCGGCAAGTTCGCTGCGCGCGCGCGCGCGCGCGCGCG 720  
DB 668 GGTGGCACTACATGCTTCCAGCGGCAAGTTCGCTGCGCGCGCGCGCGCGCGCGCGCG 727  
QY 721 TTCACTTACGACATCAAGCTGCGCAAGCAAGCACTTGGCGCATGCGCTGCTGCGCAAGCGC 780

DB 728 TTCACTTACGACATCAAGCTGCGCAAGCAAGCACTTGGCGCATGCGCGCTGCGCAAGCGC 787  
QY 781 CAATACCGGATCCGGAGCTTTGATGCGCAACCGGCAACAGATGCGCTGAGCTCC 840  
DB 788 CAATACCGGATCCGGAGCTTTGATGCGCAACCGGCAACAGATGCGCTGAGCTCC 847  
QY 841 ACCATCTGCGCTTCAAGCAGAGGTGAAGATTTGGCTGCGAGATCTTACTCAGAGCAG 900  
DB 848 ACCATCTGCGCTTCAAGCAGAGGTGAAGATTTGGCTGCGAGATCTTACTCAGAGCAG 907  
QY 901 AACGGCTCTTCTATGACCTTACTGAGCAGAGCTCTTTACGGGCTTCTATCTAT 960  
DB 908 AACGGCTCTTCTATGACCTTACTGAGCAGAGCTCTTTACGGGCTTCTATCTAT 967  
QY 961 GCGGACGAGATACCCCAAGAGATGAGATGAGCAAGCGGCGGCTCCAGGCAAGGAA 1020  
DB 968 GCGGACGAGATACCCCAAGAGATGAGATGAGCAAGCGGCGGCTCCAGGCAAGGAA 1027  
QY 1021 CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGACCCCAACTGATGAGGCGGCGG 1080  
DB 1028 CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGACCCCAACTGATGAGGCGGCGG 1087  
QY 1081 GTGAGTGAAGCGGCTTCAAGCTTCAAGGCTCACTCTCCGCTTTTTCCTTCATT 1140  
DB 1088 GTGAGTGAAGCGGCTTCAAGCTTCAAGGCTCACTCTCTGCTCTTTTTCCTTCATT 1147  
QY 1141 AAATCCAAACCTTTTATTCA 1161  
DB 1148 AAATCCAAACCTTTTATTCA 1168

RESULT 7  
BC011699  
LOCUS  
DEFINITION  
Homo sapiens C1q and tumor necrosis factor related protein 2, mRNA  
(CDNA clone MGC:19702 IMAGE:3530165), complete cds.  
ACCESSION  
BC011699  
VERSION  
BC011699.2 GI:33873400  
KEYWORDS  
MGC.  
SOURCE  
Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schlier, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,  
Scheltz, T.E., Brownstein, M.J., Ueda, T.B., Toshiyuki, S.,  
Garnica, P., Prange, C., Rabe, S.S., Loggellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahy, J., Helton, E., Kerteman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakeley, R.W., Touchman, D.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smallue, D.B.,  
Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL  
PUBMED  
2 (bases 1 to 2395)  
REFERENCE  
AUTHORS  
Strausberg, R.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (30-JUN-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,





Oy	435	CATTGGGCGGGCTGGCCCCGCTGGCCCCAGGGGGTCAACGGTACCCCGGGAAAGCATGG	494
Db	16082	CATTGGGCGGGCTGGCCCCGCTGGCCCCAGGGGGGTCAACGGTACCCCGGGAAAGCATGG	16022
Oy	495	CACACCAAGCAGAAGGGGGCCCAAGGGCAGAAAGGGGAGCAGAGCCTCCAGAGCCCTG	554
Db	16022	CACACCAAGCAGAAAGGGGGCCCAAGGGCAGAAAGGGGAGCAGAGCCTCCAGAGCCCTG	15986
Oy	555	CAGCTGTGGCAGTGGGCATACCAAGTCAAGCTTTCTCGGTGGCACTGACCAAGACTACCC	614
Db	15962	CAGCTGTGGCAGTGGGCATACCAAGTCAAGCTTTCTCGGTGGCACTGACCAAGACTACCC	15902
Oy	615	ACGGGAACGGCTGGCCCATCAAGTTTGCACAAATTCTGATGAAACGAGGTGGCCACTACAA	674
Db	15902	ACGGGAACGGCTGGCCCATCAAGTTTGCACAAATTCTGATGAAACGAGGTGGCCACTACAA	1584
Oy	675	TGCTTTCAGCGGCAGATTGCTGCTCGCGCGTGGCTCGGATCTACTACTTCACTTACGACAT	734
Db	15842	TGCTTTCAGCGGCAGATTGCTGCTCGCGCGTGGCTCGGATCTACTACTTCACTTACGACAT	1578
Oy	735	CAGCGTGGCCCAACAGCACTCTGGCCATGGCCTTGTGTGCAACGGGCCAGTACCGCATCCG	794

795 GACCTTTGATGGCAAACACCGGCAACCAAGATGTGGCCCTCAGGCTTCACCATCTCGGCTCT 854

Db 15722 GACCTTGTATGCCAACACCGGCAACACGATGTGGCCCTCAGGCTCCACCATCTGGCTCT 1566

[illegible]

DG 13882 CAGCGGGTACGATAGTGCTTCGTTCCATCAGATCA TGGG

Ov 915 TGACCCTTAATTGAAGAAGGCCTTTAAACGGGTCTTCTAATCATATGACCGAACAGGATGA 974

Db 15602 TGACCTTACTGACACAGCCTCTTACGGGCTTCTTAATCTATGCCGACGAGATGA 1554

975 CCCCAACGAGTATGACATGCCACGCGGTCTCCAGGCAAGGAAACAAGCTTCTGCACT 1034

Db 15542 CCCACACAGTATAGACATGCCACGGGGCTCTTCAGGCAGGGACAAGCTTCTGACT 1548

1035 TGGGCTTACAGACGAAGACCCCAACTTTAGGCTGGGGGTTGCGAGTGAACGGT 1094

Db 15482 TGGGCTTACAGACGACCCCAACTGTAGGCTGGGGGGTTCGAGTGCAGCCGT 1542

1095 TCTAGCCTCAGGCTCACTCTCCGCTTTTTCCTTCAATTAAATCCAACCTTT 1154

Db 15422 TCTAGCCTCAGGCTCACCTCCTCTGCTCTTTTTCCTTTTCCCTTCATTAAATCCAAACCTTT 1536

0y 1155 T7ATCA 1161  
|||||||

Db 15362 TTATICA 15356

## RESULT 10

AC112191	153248 bp	DNA	linear	PRI 23-APR-2002
LOCUS				
DEFINITION	Homo sapiens chromosome 5 clone RP11-34P1 complete sequence.			

DELINITION	AC112191.2	GI:20270112
DELINITION	AC112191.2	GI:20270112
ACCESSION	AC112191	
ACCESSION	AC112191.2	GI:20270112

KEYWORDS	HTG.
SOURCE	Homo sapiens (human)

ORGANISM      Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;

REFERENCE  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 153248)  
DOI: 10.1002/ajp.10001 and GenBank Human Genome Center  
Accession Number: U00001

AUTHORS  
 TITLE  
 Direct Submission  
 Journal, Published

**ORIGINAL SOURCE**  
2 (bases 1 to 153248)  
**REFERENCE**  
DOE Joint Genome Institute.  
**AUTHORS**

TITLE	Direct Submission
JOURNAL	Submitted (20-FEB-2002)
	Production Sequencing Facility, DOE Joint

REFERENCE 3 (bases 1 to 153248)  
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-APR-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 Db On Apr 23, 2002 this sequence version replaced gi:18767463.  
 COMMENT Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 www.ehgc.stanford.edu  
 Quality: Phrap Quality >=40 99.9% of Sequence;  
 Estimated total Number of Errors is 0.1.  
 NOTE: This insert is not the entire sequence of the clone (entire sequence is 171.7kb). It is clipped at the overlap with AC008720.  
 The number of bases overlapped is 36260.

FEATURES  
 source  
 location/Qualifiers  
 1..153248  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="RP11-34P1"

ORIGIN  
 Query Match 67.5%; Score 783.8; DB 9; Length 153248;  
 Best Local Similarity 99.7%; Pred. No. 6,16-141;  
 Matches 785; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 375 AGGTTCACCTGCGCGGACGTAACCGGGAAGCCAGACCAAAAGGCGAAAGCCGGGCGC 434
Db 20117 AGGTTCACCTGCGCGGACGTAACCGGGAAGCCAGACCAAAAGGCGAAAGCCGGGCGC 20176
Qy 435 CATTGGCGGGCTGCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 494
Db 20177 CATTGGCGGGCTGCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20236
Qy 495 CACACGAGCAAGAGAGGGGCGCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 554
Db 20237 CACACGAGCAAGAGAGGGGCGCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20296
Qy 555 CAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 614
Db 20297 CAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20356
Qy 615 ACGGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 674
Db 20357 ACGGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20416
Qy 675 TGCTTCAGCGGCGGCAAGTTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 734
Db 20417 TGCTTCAGCGGCGGCAAGTTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20476
Qy 735 CAGCGTGGCGCAAGAGCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 794
Db 20477 CAGCGTGGCGCAAGAGCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20536
Qy 795 GACCTTGATGCGCAACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 854
Db 20537 GACCTTGATGCGCAACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20596
Qy 855 CAAGCAGAGGTGACGAGTTGGCTGCGAGATCTTCTACTCAGAGCAGAAAGCGGCGCTCTTCTA 914
Db 20597 CAAGCAGAGGTGACGAGTTGGCTGCGAGATCTTCTACTCAGAGCAGAAAGCGGCGCTCTTCTA 20656
Qy 915 TGACCCCTTACCTGACACAGCTCTTTACGGGCGTCTCTTAATCTATGCGGCGGCGGCGGCGG 974
Db 20657 TGACCCCTTACCTGACACAGCTCTTTACGGGCGTCTCTTAATCTATGCGGCGGCGGCGGCGG 20716
Qy 975 CCCCAAGAGGTATAGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1034
Db 20717 CCCCAAGAGGTATAGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20776
Qy 1035 TGGGCTTACAGAGCAGAACCCCAACACTGTAAGGCTGGGGGCTGGGGGCTGAGTGAAGCGGT 1094

```

Db 20777 TGGGCTTACAGAGCAGAACCCCAACACTGTAAGGCTGGGGGCTGGGGGCTGAGTGAAGCGGT 20836

Qy 1095 TCTAGCTCAGGCTCAGCTCTCTCGGCTCTTTTTCCTTCTATTAATCAACCACTTT 1154

Db 20837 TCTAGCTCAGGCTCAGCTCTCTCGGCTCTTTTTCCTTCTATTAATCAACCACTTT 20896

Qy 1155 TTATTCAC 1161

Db 20897 TTATTCAC 20903

RESULT 11  
 BC030324  
 LOCUS  
 DEFINITION Mus musculus C1q and tumor necrosis factor related protein 2, mRNA  
 (CDNA clone MGC:40933 IMAGE:5375090), complete cds.  
 ACCESSION BC030324  
 VERSION BC030324.1 GI:20988949  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Kjauner R.D., Collins E., Feingold E.A., Grouse L.H., Derge J.G., Strausberg R.L., Wagner F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haileh F., Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheet T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hu X., Gibbs R.A., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kesteman M., Madan A., Young A.C., Shevchenko Y., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.B., Jones S.J. and Marra M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL  
 PUBMED 12477932  
 REFERENCE 2 (bases 1 to 1236)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)  
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Ahlter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Khong, P., Latic, P., Legaspi, R., Madu, O.L., Masello, C., Maskeri, B., Mastrion, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stancirpop, S., Thomas, P.J., Touchman, J.W., Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
Series: IRAX Plate: 65 Row: m Column: 7  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES  
source

Location/Qualifiers  
1..1236  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N-3"  
/db\_xref="taxon:10090"  
/clone="MGC:40933 IMAGE:5375090"  
/issue\_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."  
/clone\_1db="NCI CGAP\_Mam2"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..1236  
/gene="C1qtnf2"  
/db\_xref="LocusID:69183"  
/db\_xref="MGI:1916433"  
140..1024  
/gene="C1qtnf2"  
/codon\_start=1  
/product="C1qtnf2 protein"  
/protein\_id="AAH0324.1"  
/db\_xref="GI:20988950"  
/db\_xref="LocusID:69183"  
/db\_xref="MGI:1916433"  
/translation="MTIFKRYTMTISVTLALCALPCAADPMLGAFARDDPKGKPOLVCSLPGPDPPEAGSGSVGKMGPGDGDGDGDSDSDSEBPGRGKQGPGRKATGRAGPRGKGVSGTPGKGTGKPKGPKGKBPGLPGKSGSSKAKSAFSVAIKSYPRERLPIKFDKILMKGHNASGKFCVSGIYFTYDITLANKHIALGVHNGQYRIPTFDANTGNHDVASGSTIILAKGDEVMLQIFYBQNGHFPYPMYTD SLFTGFLIVADQDPNEV"  
599..1000  
/gene="C1qtnf2"  
/note="C1q Region: Complement component C1q domain. Globular domain found in many collagens and epouonously in complement C1q. When part of full length proteins these domains form a 'bouquet' due to the multimerization of heterotrimers. The C1q fold is similar to that of tumour necrosis factor"  
/db\_xref="CDD:smat00110"

## misc\_feature

599..1000  
/gene="C1qtnf2"  
/note="C1q Region: Complement component C1q domain. Globular domain found in many collagens and epouonously in complement C1q. When part of full length proteins these domains form a 'bouquet' due to the multimerization of heterotrimers. The C1q fold is similar to that of tumour necrosis factor"  
/db\_xref="CDD:smat00110"

## ORIGIN

Query Match 62.0%; Score 719.4; DB 10; Length 1236;  
Best Local Similarity 82.2%; Pred. No. 2.7e-128;  
Matches 851; Conservative 0; Mismatches 181; Indels 3; Gaps 2;

QY 121 AGGGTAACACATGATCCCTGGGCTCTCTGGCTGTGCCCTCCCTGTGCTGTAC 180  
DB 155 AAGGTGACACATGATCTCTGGGCTCTCTGGCTGTGCCCTCCCTGTGCTGTAC 214  
QY 181 CCACTGCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 240  
DB 215 CCAATGCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 274  
QY 241 AGCTGCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 300  
DB 275 AGCTGCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 334  
QY 301 ATGGAGCAATGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 360  
DB 335 GTGGAGCAATGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 394  
QY 361 GACAGCGAGAGAGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 420  
DB 395 GACAGCTGAGAGAGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 454  
QY 421 GGCAAGCGCGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGG 480

DB 455 GGCAAGCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGG 514  
QY 481 CCCGGAGAGATGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 540  
DB 515 CCCGGAGAGATGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 574  
QY 541 CTCCTGAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 600  
DB 575 CTCCTGAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 634  
QY 601 ACCAAGAGCTACCAAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 660  
DB 635 ACCAAGAGCTACCAAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 694  
QY 661 GGTGGCACTACCAAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 720  
DB 695 GGTGGCACTACCAAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 754  
QY 721 TTACCTGAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 780  
DB 755 TTACCTGAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 814  
QY 781 CAGTACCGCATCCGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 840  
DB 815 CAGTACCGCATCCGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 874  
QY 841 ACATCTGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 900  
DB 875 ACATCTGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 934  
QY 901 AACGGCTCTCTGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGG 960  
DB 935 AATGGCTCTCTGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGG 994  
QY 961 GCCAGCAGATGACCCCAAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 1019  
DB 995 GCTGACCAAGAGACCCCAAGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 1054  
QY 1020 ACAAGCTCTGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGG 1079  
DB 1055 CTAAGAGCTCTG--GCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 1112  
QY 1080 GCTGAGTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 1139  
DB 1113 AATCTGGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCT 1172  
QY 1140 TAAATCAAGCTCT 1154  
DB 1173 TAAATCAAGCTATT 1187

RESULT 12  
BD209667 1107 bp DNA linear PAT 17-JUN-2003  
LOCUS BD209667  
DEFINITION Compositions isolated from skin cells and methods for their use.  
ACCESSION BD209667.1 GI:33019437  
VERSION BD209667.1  
KEYWORDS JP 2002512798-A/139.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 1107)  
Strachan,L., Sleeman,M., Watson,J.D., Onrust,R., Kumble,A. and Murison,J.G.  
TITLE Compositions isolated from skin cells and methods for their use  
JOURNAL Patent: JP 2002512798-A 139 08-MAY-2002;  
GENESIS RESEARCH AND DEVELOPMENT CORP LTD  
OS Rattus sp. (rat)  
PN JP 2002512798-A/139  
PD 08-MAY-2002



Db 582 CTCACAGGCCCCCTGAGCTGGGCGAGTAGCCGACCAAGTCGACCTTTTCGTTGGCGGTGA 641  
Qy 601 ACCAAGAGCTACCCACGCGGAGCGGCTGCCCATCAAGTTTGACAGATTCTGATGAACGAG 660  
Db 642 ACCAAGAGTTACCCACGAGCGAGCTGCCCATCAAGTTTGACAGATTCTGATGAATAG 701  
Qy 661 GGTGGCCCACTCAATAGCTTCCAGCGGCGCAAGTTCCGCTGCGGCGCTGGGATCTACTAC 720  
Db 702 GGAAGCCCACTCAATAGCTCAAGTAGGCAAGTTTCGCTGCAAGCGGCCAGGAGTTCTATTAC 761  
Qy 721 TTCACTTACGACATCAAGCTGGCCCAACAGCACTGGCCATCGGCTGAGTGCACACCGGC 780  
Db 762 TTTCATATGACATTAGCTGGCCCAACACCTGGCCATCGGCTGAGTGCACAGATGGC 821  
Qy 781 CAGTACCGCATCCGAGCTTTGATGCCAAACCGGCAACACAGATGAGGCTCAGGCTCC 840  
Db 822 CAGTACCGCATTCGAGCTTTGAGCGCCCAACCGGCAACACAGATGAGGCTCAGGCTCC 881  
Qy 841 ACATCTCGGCTCTCAAGCAGAGGAGGAGTTTGGCTGAGAGTCTTACTCAGAGCAG 900  
Db 882 ACATCTCGGCTCTCAAGCAGAGGAGGAGTGTGTTTACAGATTCTTACTCGAGCAG 941  
Qy 901 AACGGGCTCTTCTATGACCTTACTGACAGACAGCTCTTTACGAGGCTTCTATATCTAT 960  
Db 942 AATGAGCTCTTCTAAGACCTTATTTAGACGACAGCTTCTACCGGCTTCTCATCTAC 1001  
Qy 961 GCCACACAGATGACCCCAACAGAGTATGACATGCCACGG--CGGTCTTCAGGCGAGG 1018  
Db 1002 GCTGATCAAGAGAGCCCCAATGAGTATGACAGAGCTGGGCTTGAAGCCGTCAGGCGAGG 1061  
Qy 1019 AACAGGCTTCTGAGCTTGGGCTTACAGAGCAAGACCCCAACATG 1063  
Db 1062 ACTAAGATTCCGCAAGGCTGCTGATGAAGAGATCTTGAACTG 1106

RESULT 14  
BD209734 1052 bp DNA linear PAT 17-JUL-2003  
LOCUS BD209734  
DEFINITION Compositions isolated from skin cells and methods for their use.  
ACCESSION BD209734  
VERSION BD209734.1 GI:33019504  
KEYWORDS JP 2002512798-A/206.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1052)  
Strachan,L., Sleeman,M., Watson,J.D., Onrust,R., Kumble,A. and  
Murison,J.G.  
Compositions isolated from skin cells and methods for their use  
Patent: JP 2002512798-A 206 08-MAY-2002;  
GENESIS RESEARCH AND DEVELOPMENT CORP LTD  
OS Rattus sp. (rat)  
PN JP 2002512798-A/206  
PD 08-MAY-2002  
PF 29-APR-1999 JP 2000546009  
PR 29-APR-1998 US 09/069726, 09-NOV-1998 US 09/188930 PT  
LORNA STRACHAN, MATTHEW SLEEMAN, JAMES DOUGLAS WATSON, RENE PI  
ONRUST,  
PI ANAND KUMBLE, JAMES GREG MURISON  
PC C12N15/09, A61K38/00, A61P9/00, A61P17/00, A61P29/00, A61P31/18, PC  
A61P35/00,  
PC 07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12N15/ PC  
00,  
PC A61K37/02, C12N5/00  
CC Compositions isolated from skin cells and methods for their  
CC use.  
FT Key Location/Qualifiers  
FT source 1..1052  
FT /organism='Rattus sp. (rat)'.  
FEATURES  
Location/Qualifiers

source 1..1052  
/organism='Rattus sp.'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:10118'  
ORIGIN  
Query Match 60.4%; Score 701.6; DB 6; Length 1052;  
Best Local Similarity 82.9%; Pred. No. 7.4e-125;  
Matches 837; Conservative 0; Mismatches 169; Indels 4; Gaps 3;  
Qy 121 AGGGTAACCAATGATCCCTGAGTGTCTCTGAGCTGTGCTCCCTGTGCTGTGAC 180  
Db 34 AAGGTGACCAATGATCTCTGATGTCTTGGCTGTGCTCTTCCGTGTGCTGTGAC 93  
Qy 181 CCACTGCTGGCGCTTTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAACTGTCTGC 240  
Db 94 CCAATGCTGTGTGCTTGTGCTGCGAGGAGCTTCCAGAAAGGCTGTCTCACTGTGTGC 153  
Qy 241 AGCTGCTGCGCCCGAGGCGCCACCGGCGCCCGAGGCGCCCGAGGCGCTCAGGAATG 300  
Db 154 AGTGGCTGTGCTCCCAAG--CACTGCGCTTCCAGAGCAACAGATCTCAGGAATG 212  
Qy 301 ATGGACGAATGAGCTTCTCTGCGAAAGCGCCAGATGACACAGCGGACCGGAGG 360  
Db 213 GTGGGAAGATGGGTTTCTGTGATGAGTGGCCAGACCGCCAGAGCGAGACCGAGG 272  
Qy 361 GACACGGAGAGAGTGTCCACTGTGCGGACAGGTAACTGGGGAAGCCAGACCAAG 420  
Db 273 GACAGTGAAGAGAGGTCCACCTGCGAGGACAGGCAACAGGAAACAGAGACCAAG 332  
Qy 421 GGCAAGCGGGGCGCATTTGGCGGGCTGAGCCCGGCGCCCGAGGCGGTACCGGTAC 480  
Db 333 GGCAAGCTGGGCGCATTTGGAGAGCGGCTCTGAGAGCCCAAGGGGCTAGGTGTAAC 392  
Qy 481 CCCGGAAGCATGCAACCAAGGAGAGGCGCCCAAGGCAAGAAAGGAGCCAGGC 540  
Db 393 CCCGGAAGCATGCTATACCGGGCAAGAGGCACTTAAGGCAAGAAAGGAGACCTGGG 452  
Qy 541 CTCCAGGCGCTTCAAGCTGTGCGAGTGGCCATACAGTCAAGTCTTCTCGGTGCAATG 600  
Db 453 CTCCAGGCGCTTCAAGCTGTGCGAGTGGCCAGTCAAGTCTTCTCGGTGCAATG 512  
Qy 601 ACCAAGAGCTACCCACGAGGAGCGGCTGCGCCATCAAGTTTGAAGATTTCTGATGACGAG 660  
Db 513 ACCAAGAGTTACCCACGAGGAGCGGCTGCGCCATCAAGTTTGAAGATTTCTGATGATGAG 572  
Qy 661 GGTGGCCCACTCAATAGCTTCCAGCGGCAAGTTTCGTGCGGCTGAGCTGGATCTACTAC 720  
Db 573 GGAAGCCCACTCAATAGCTACAGTAGGCAAGTTTCGTGCGGCTGAGCTGGATCTACTAC 632  
Qy 721 TTCACTTACGACATCAAGCTGGCCCAACAGCACTGGCCATCGGCTGTGTGACACGCGC 780  
Db 633 TTTCATATGACATTAGCTGGCCCAACCACTGGCCATCGGCTGTGTGACACAGTGC 692  
Qy 781 CAGTACCGCATCCGAGCTTTGATGACCAACCGGCAACAGATGAGGCTCAGGCTCC 840  
Db 693 CAGTACCGCATTCGAGCTTTGACGCAACCGGCAACAGATGAGGCTCAGGCTCC 752  
Qy 841 ACATCTCGGCTCTCAAGCAGAGGAGTGTGCTGACAGATCTTCTACTCAGAGCAG 900  
Db 753 ACATCTCGGCTCTCAAGCAGAGGAGTGTGCTGACAGATCTTCTACTCAGAGCAG 812  
Qy 901 AACGGGCTCTTCTATGACCTTACTGACAGACAGCTCTTTACGGGCTTCTATATCTAT 960  
Db 813 AATGAGCTCTTCTAAGACCTTATTTAGACGACAGCTGTTCACCGGCTTCTCATCTAC 872  
Qy 961 GCCACAGAGATGACCCCAACAGAGTATGACAGTATGAGCCAGCG--CGGTCTTCCAGGCGAG 1019  
Db 873 GCTGATCAAGAGAGCCCAATGAGTATGACAGAGCTGGGTTAGAGTCCAGGACAGGA 932  
Qy 1020 AACAGCTTCTGAGCTTGGGCTTACAGAGCAAGACCCCAACCTGTAGGCTGGGAGTGGG 1079  
Db 933 CTAAAGATTCCGCAAGGCTGTGATGAAGAGAGTCTGAACCTGAGCTGGGAGCTGGCA 992



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2004, 02:28:47, Search time 655 Seconds

(without alignments)  
9304.699 Million cell updates/sec

Title: US-10-621-787-1

Perfect score: 1161  
Sequence: 1 ggaataactagctctggggcc.....aatccaactttattca 1161

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapect 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1161	100.0	1161	3 AAC65132	AAC65132 Human adi
2	1161	100.0	1161	8 ABZ24307	ABZ24307 Human GMG
3	1161	100.0	1161	11 ADM66421	ADM66421 CDNA enco
4	1161	100.0	1161	12 ADL26856	ADL26856 Human zac
5	1154.2	99.4	1171	4 AAS31128	AAS31128 Human dia
6	1148.4	98.9	1178	8 ACD03627	ACD03627 Novel hum
7	1140.4	98.2	1178	12 ADM93810	ADM93810 CDNA enco
8	1120.8	96.5	1171	6 ABO54564	ABO54564 Human ova
9	1084	93.4	1211	10 ADF94944	ADF94944 Human gen
10	1082.8	93.3	1114	10 ADF94916	ADF94916 Human gen
11	988.4	85.1	993	8 ACC42468	ACC42468 Human ClQ
12	930.4	80.1	932	3 AAZ45606	AAZ45606 CDNA enco
13	703.6	60.6	885	8 ACC42469	ACC42469 Murine CI
14	702.2	60.5	1107	3 AAZ61744	AAZ61744 CDNA enco
15	702.2	60.5	1107	4 AAC99677	AAC99677 Skin cell
16	702.2	60.5	1107	6 ABL34829	ABL34829 Rat CDNA
17	701.6	60.4	1052	3 AAZ61811	AAZ61811 CDNA enco
18	701.6	60.4	1052	4 AAC99744	AAC99744 Skin cell
19	695	59.9	855	3 ABL34896	ABL34896 Rat CDNA
20	627.4	54.0	886	3 AAC67788	AAC67788 Murine AC
21	627.4	54.0	886	3 AAZ45688	AAZ45688 CDNA enco

22 627.4 54.0 886 8 ABZ24308  
23 613.6 52.9 855 3 AAC65137  
24 613.6 52.9 855 11 ADM66430  
25 613.6 52.9 855 12 ADL26865  
26 507 43.7 855 3 AAZ45689  
27 507 43.7 855 3 AAZ45607  
28 453 39.0 750 3 AAZ61624  
29 453 39.0 750 4 AAC99557  
30 453 39.0 750 6 ABL34709  
31 337.4 29.1 909 4 AAC89872  
32 309 26.6 1282 4 AAC89875  
33 295.4 25.4 870 4 AAD16350  
34 295.4 25.4 870 6 ABO86182  
35 295.4 25.4 870 10 ADF8251  
36 295.4 25.4 912 4 AAC89867  
37 295.4 25.4 912 4 AAD16351  
38 295.4 25.4 1242 6 AAL44067  
39 295.4 25.4 1297 6 ABK94966  
40 295.4 25.4 1297 10 ADF8249  
41 295.4 25.4 1304 8 ADA40160  
42 295.4 25.4 1304 10 ADC73711  
43 295.4 25.4 1304 10 ADD37685  
44 295.4 25.4 1304 10 ADA56322  
45 295.4 25.4 1336 10 ADF94915

## ALIGNMENTS

RESULT 1  
ID AAC65132 standard; CDNA; 1161 BP.  
XX  
AC AAC65132;  
XX  
DT 12-FEB-2001 (first entry)  
XX  
DE Human adipocyte complement related protein homologue zacr2p2 CDNA.  
XX  
KW Human; zacr2p2; adipocyte complement related protein; Acrp30;  
XX energy balance; metabolism; haemostasis; anti-microbial; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200063376-A1.  
XX  
PD 26-OCT-2000.  
XX  
PE 19-APR-2000; 2000WO-US010452.  
XX  
PR 20-APR-1999; 99US-00295072.  
XX  
XX (ZYMO ) ZYMOGENETICS INC.  
PA Piddington CS, Bishop PD;  
XX WPI; 2000-647517/62.  
XX P-PSDB; AAB30232.  
XX  
PT Human DNA sequence encoding a zacr2p2 polypeptide which has homology to an  
PT adipocyte complement related protein (acrp30), useful in gene therapy  
PT applications for inhibiting or increasing zacr2p2 activity.  
XX  
XX Claim 31; Page 110-112; 125pp; English.  
XX  
XX The present invention is related to the isolation and uses of a homologue  
XX to the adipocyte complement related protein Acrp30, known as zacr2p2. The  
XX zacr2p2 protein is involved in energy balance, and the protein, its  
XX antibodies and coding sequence can be used to modulate energy balance,  
XX haemostasis, calcium ion concentration, muscle contraction, hormone  
XX secretion, DNA synthesis and cell growth, inositol phosphate turnover,  
XX archidonate release, phospholipase-C activation, gastric emptying, human  
XX neutrophil activation or ADCC capability and superoxide anion production.







Query Match 100.0%; Score 1161; DB 11; Length 1161;  
 Best Local Similarity 100.0%; Pred. No. 7e-263;  
 Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GAAAAATATATGCTTGGGCGGACGCTCTGCGGCTGCTGCGCTGAGAGAAAGCGGAGAC 60
DB 1 GAAAAATATATGCTTGGGCGGACGCTCTGCGGCTGCTGCGCTGAGAGAAAGCGGAGAC 60
QY 61 GCGGAGCGGCGGCGGAGCTTCTTGTCCGGAAGCCCTGGAAGTGGCGGCGGCGGCGG 120
DB 61 GCGGAGCGGCGGCGGAGCTTCTTGTCCGGAAGCCCTGGAAGTGGCGGCGGCGGCGG 120
QY 121 AGGGTAACACATGATCCCTGAGGTGCTCTGAGCTGTGCGCTCCCTGTGCTGCTGAC 180
DB 121 AGGGTAACACATGATCCCTGAGGTGCTCTGAGCTGTGCGCTCCCTGTGCTGCTGAC 180
QY 181 CCACTGCTGGCGCTTGTGCTCGAGGGAATTCCGAAAGGCTCCCTCAACTGCTGCTGC 240
DB 181 CCACTGCTGGCGCTTGTGCTCGAGGGAATTCCGAAAGGCTCCCTCAACTGCTGCTGC 240
QY 241 AGCCTGCTGGCGCTTGTGCTCGAGGGAATTCCGAAAGGCTCCCTCAACTGCTGCTGC 300
DB 241 AGCCTGCTGGCGCTTGTGCTCGAGGGAATTCCGAAAGGCTCCCTCAACTGCTGCTGC 300
QY 301 ATGGGAGCAATGGGCTTCTGCGAAAGACGCGCAAGATGACACAGCGGCGGCGG 360
DB 301 ATGGGAGCAATGGGCTTCTGCGAAAGACGCGCAAGATGACACAGCGGCGGCGG 360
QY 361 GACAGCGGAGAGGAGGTCCACTGTGCGGACGAGTAACCGGGGAAAGCGAGCAAAAG 420
DB 361 GACAGCGGAGAGGAGGTCCACTGTGCGGACGAGTAACCGGGGAAAGCGAGCAAAAG 420
QY 421 GCGAAAGCGGCGGCGCTTGGGCGGCGCTGCGCGCTGCGCGCTCAAGGAGTCAAGCTAAC 480
DB 421 GCGAAAGCGGCGGCGCTTGGGCGGCGCTGCGCGCTGCGCGCTCAAGGAGTCAAGCTAAC 480
QY 481 CCGGAGAGATGGCACACGAGCAAGAGGCGCCCAAGGCGAAGAAAGGAGGCCAGGC 540
DB 481 CCGGAGAGATGGCACACGAGCAAGAGGCGCCCAAGGCGAAGAAAGGAGGCCAGGC 540
QY 541 CTCCCAAGGCGCTGCGAGCTGTGCGCAATCAAGTCAAGTCTCTGCTGCGAGTGC 600
DB 541 CTCCCAAGGCGCTGCGAGCTGTGCGCAATCAAGTCAAGTCTCTGCTGCGAGTGC 600
QY 601 ACCAAGAGCTTACCAACGAGGCGGCTGCCCATCAAGTTTGAACAAGTTTGAATGAAGAG 660
DB 601 ACCAAGAGCTTACCAACGAGGCGGCTGCCCATCAAGTTTGAACAAGTTTGAATGAAGAG 660
QY 661 GGTGGCACTTACATGCTTCCAGGCGGCAAGTTCGTGCGGCGTGCCTGGGATCTACTAC 720
DB 661 GGTGGCACTTACATGCTTCCAGGCGGCAAGTTCGTGCGGCGTGCCTGGGATCTACTAC 720
QY 721 TTCACTTACATGATCAGCTGCTGCAACAGCACTGCGCATGCGCTGTGTGCAACGCGC 780
DB 721 TTCACTTACATGATCAGCTGCTGCAACAGCACTGCGCATGCGCTGTGTGCAACGCGC 780
QY 781 CAGTACCGCAATCCGGAAGCTTGTGATGCAACACGCGCAACAGATGTGGCTCAGGCTCC 840
DB 781 CAGTACCGCAATCCGGAAGCTTGTGATGCAACACGCGCAACAGATGTGGCTCAGGCTCC 840
QY 841 ACCATCTGAGCTCTCAAGACGAGGTGAGCAAGTTTGGCTGCGAGATCTTCTACTCAGACAG 900
DB 841 ACCATCTGAGCTCTCAAGACGAGGTGAGCAAGTTTGGCTGCGAGATCTTCTACTCAGACAG 900
QY 901 AACGGGCTCTTATATGACCTTACTGAGACAGCTCTTTTACGGGCTTCTTATCTAT 960
DB 901 AACGGGCTCTTATATGACCTTACTGAGACAGCTCTTTTACGGGCTTCTTATCTAT 960
QY 961 GCGGACCAAGATGACCCCAACGAGGTATGACATGCGCGGCTCTCCAGGAGGAGAA 1020
DB 961 GCGGACCAAGATGACCCCAACGAGGTATGACATGCGCGGCTCTCCAGGAGGAGAA 1020

```

```

QY 1021 CAAGCTTCTGAGCTTGGGCTTATAGAGCAAGACCCCAACTGATAGGCTGGGCTGGGCG 1080
DB 1021 CAAGCTTCTGAGCTTGGGCTTATAGAGCAAGACCCCAACTGATAGGCTGGGCTGGGCG 1080
QY 1081 GTGAGTGTAGCGGCTTCTAGGCTTCAAGGCTCACTCTCGGCTCTTTTTCCTTCAATT 1140
DB 1081 GTGAGTGTAGCGGCTTCTAGGCTTCAAGGCTCACTCTCGGCTCTTTTTCCTTCAATT 1140
QY 1141 AAATCCAAACCTTTTATTTCA 1161
DB 1141 AAATCCAAACCTTTTATTTCA 1161

```

## RESULT 4

```

ADL26856
ID ADL26856 standard; DNA, 1161 BP.
XX
AC ADL26856;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human zacr2p DNA.
XX
XX Adipocyte complement related protein homologue; zacr2p; energy balance;
XX energy efficiency; infection; cellular metabolism; tumour;
XX cardiopulmonary bypass ischaemia; recesitation; myocardial infarction;
XX post-trauma vasospasm; stroke; percutaneous transluminal angioplasty;
XX vasculature disease; disseminated intravascular coagulation;
XX non-vascular disease; arteriosclerosis; prosthetic biomaterial;
XX surgical equipment; wound healing; gene therapy; human; chromosome 5;
XX gene; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 133..990
XX FT /*tag= b
XX FT /product= "zacr2p protein"
XX FT sig_peptide 133..177
XX FT /*tag= a
XX FT mat_peptide 178..987
XX FT /*tag= c
XX FT /product= "Mature zacr2p protein"
XX
XX US2004024187-A1.
XX
XX 05-FEB-2004.
XX
XX 17-JUL-2003; 2003US-00621787.
XX
XX 20-APR-1999; 99US-0130207P.
XX 19-APR-2000; 2000US-00552204.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Piddington CS, Bishop PD;
XX
XX WPI: 2004-142676/14.
XX DR P-PSDB; ADL26857.
XX
XX New zacr2p polypeptide, useful in modulating energy balance, treating
XX tumour, myocardial infarction, stroke, disseminated intravascular
XX coagulation or arteriosclerosis, mediating wound repair and healing or in
XX enhancing antibody formation.
XX
XX Claim 31; SEQ ID NO 1; 41bp; English.
XX
XX The invention relates to novel adipocyte complement related protein
XX homologue, zacr2p polypeptides and polymucleotides. Sequences of the
XX invention are useful in modulating energy balance, analysing energy
XX efficiency of a mammal and in studying insulin, exogenous microorganism
XX infection, cellular metabolism, dimerisation or oligomerisation. They are
XX also useful in treating tumours, cardiopulmonary bypass ischaemia and

```

CC resection), myocardial infarction and post-trauma vasospasm e.g. stroke  
CC or percutaneous transluminal angioplasty, vascular disease e.g.  
CC attenuated intravascular coagulation and non-vascular disease e.g.  
CC atherosclerosis, zaccp2 sequences are also useful in pacifying damaged  
CC collagenous tissue, prosthetic biomaterials and surgical equipments,  
CC preventing infection, mediating wound repair, or wound healing and in  
CC enhancing antibody formation. The invention is also used in gene therapy.  
CC The present sequence is human zaccp2 DNA. zaccp2 gene is located on  
CC chromosome 5.

**SQ Sequence 1161 BP; 237 A; 375 C; 350 G; 199 T; 0 U; 0 Other;**

Query Match	Score	DB	Length
100.0%	1161	12	1161
100.0%	70	303	

Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGAAAACTATATGCTTGGGGGCGGACGCTCTGCGCTCCGGCTGTGGCCGCTGAGAGAAAGCCGGGAC	60
Ds	1	GGAAAACTATATGCTTGGGGGCGGACGCTCTGCGCTCCGGCTGTGGCCGCTGAGAGAAAGCCGGGAC	60
Qy	61	GGGAGACCCCGCGCGAGAGACTTCTTTTGTCTTCGGACGCGCCCTTGAGAGTGAGCGGCGACGCGG	120
Ds	61	GGGAGACCCCGCGCGAGAGACTTCTTTTGTCTTCGGACGCGCCCTTGAGAGTGAGCGGCGACGCGG	120
Qy	121	AGGGTAAACAACATATGATCCCTTGGGGTGTCTTGGCTGTGGCTTCCCTGTGTCTTGAC	180
Ds	121	AGGGTAAACAACATATGATCCCTTGGGGTGTCTTGGCTGTGGCTTCCCTGTGTCTTGAC	180
Qy	181	CGACTGCTTGGGGGCTTTTGTCTGCAAGGAGCTTCGGAGAAAGGCTCCGCTCAAGTGTGTC	240
Ds	181	CGACTGCTTGGGGGCTTTTGTCTGCAAGGAGCTTCGGAGAAAGGCTCCGCTCAAGTGTGTC	240
Qy	241	AGCCTGCTTGGCCCCCGAGGGGCCACCTCGGGCCCCCGAGAGGCCCAAGGGGCTCTCAGAAATG	300
Ds	241	AGCCTGCTTGGCCCCCGAGGGGCCACCTCGGGCCCCCGAGAGGCCCAAGGGGCTCTCAGAAATG	300
Qy	301	ATGGGAGCAATGGGGCTTTTCTGTGGCAAAACCGGCGCAAGATGAGACACAGCGGACACGGGGG	360
Ds	301	ATGGGAGCAATGGGGCTTTTCTGTGGCAAAACCGGCGCAAGATGAGACACAGCGGACACGGGGG	360
Qy	361	GACAGGAGAGAGAGAGGTGCACCTTGGCGCGACAGGGTAAACCGGGAGAAAGCCAGACCAAAAG	420
Ds	361	GACAGGAGAGAGAGAGGTGCACCTTGGCGCGACAGGGTAAACCGGGAGAAAGCCAGACCAAAAG	420
Qy	421	GGCAATGCCGGGGCCATTGGGGCGGGCTGTGGCCCCGTGGCCCCCAAGGGGGTCAACGGTAAAC	480
Ds	421	GGCAATGCCGGGGCCATTGGGGCGGGCTGTGGCCCCGTGGCCCCCAAGGGGGTCAACGGTAAAC	480
Qy	481	CCGGGAGAGATAGGCACACACAGGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGACCAAGGC	540
Ds	481	CCGGGAGAGATAGGCACACACAGGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGACCAAGGC	540
Qy	541	CTCCAGAGCCCTGTGAGCTGTGGAGTGGCCATACCAAGTCAAGCTTCTCGTGTGCAGTG	600
Ds	541	CTCCAGAGCCCTGTGAGCTGTGGAGTGGCCATACCAAGTCAAGCTTCTCGTGTGCAGTG	600
Qy	601	ACCAAGAGTACCCACCGGGAGGGCGCTGCCATCAAGTTTGAACAAGATTCTGATGAAAGAG	660
Ds	601	ACCAAGAGTACCCACCGGGAGGGCGCTGCCATCAAGTTTGAACAAGATTCTGATGAAAGAG	660
Qy	661	GGTGGCACTAATATGCTTTCAGCGGCAAGTTGTGTGCGGGCGTGGCTGGAGTTACTATAC	720
Ds	661	GGTGGCACTAATATGCTTTCAGCGGCAAGTTGTGTGCGGGCGTGGAGTTACTATAC	720
Qy	721	TTTCACTTAAGACATACAGCTGTGGCCAAACAAGCACTGGCGCATTCGGCTGTGGTGCACAAAGGC	780
Ds	721	TTTCACTTAAGACATACAGCTGTGGCCAAACAAGCACTGGCGCATTCGGCTGTGGTGCACAAAGGC	780
Qy	781	CAGTACCGGATCCGGACCTTTGATGTGCACACCGGCAACAGATGTGTGCTTCAAGGCTTC	840
Ds	781	CAGTACCGGATCCGGACCTTTGATGTGCACACCGGCAACAGATGTGTGCTTCAAGGCTTC	840
Qy	841	ACCATCTGGCTCTCAAGCAGGGTACGAAGTTTGGCTGACGATCTTCTACTCAGACAG	900

Accession	Gene	Chromosome	Position (kb)	Sequence	Length (bp)
D8	ACCATCTGGGTCTCAAGCAGAGGGTGAAGAGTTTGGCTGAGATCTTCTACTCAGAGCAG	900			
QY	901 AACGGGCTCTTCTATGACCTTACTGACACAGACAGCTCTTTACGGGCTTCTATCTAT	960			
Db	901 AACGGGCTCTTCTATGACCTTACTGACACAGACAGCTCTTTACGGGCTTCTATCTAT	960			
QY	961 GCCGACACAGATGACCCCAACGAGGTATATGACATGACGAGGGGCTCTCCAGAGGAGAA	1020			
Db	961 GCCGACACAGATGACCCCAACGAGGTATATGACATGACGAGGGGCTCTCCAGAGGAGAA	1020			
QY	1021 CAAGCTTCTGACCTTGGGCTTACAGACAGAACCCCAAACTGATGGCTGGGGTGGGGG	1080			
Db	1021 CAAGCTTCTGACCTTGGGCTTACAGACAGAACCCCAAACTGATGGCTGGGGTGGGGG	1080			
QY	1081 GTGAGTGAAGCGGTTCTAGCCTCAGGCTCACCTCTCCGCTCTTTTTCCTTCAAT	1140			
Db	1081 GTGAGTGAAGCGGTTCTAGCCTCAGGCTCACCTCTCCGCTCTTTTTCCTTCAAT	1140			
QY	1141 AATCCAACTCTTTTATATCA	1161			
Db	1141 AATCCAACTCTTTTATATCA	1161			
RESULT 5					
AAS31128					
ID	AAS31128 standard; cDNA; 1171 BP.				
AC	AAS31128;				
XX					
DT	04-DEC-2001 (first entry)				
XX					
DE	Human diagnostic and therapeutic polynucleotide (DTHP) #143.				
XX					
KW	Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;				
KW	cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;				
KW	acquired immune deficiency syndrome; AIDS; autoimmune disorder;				
KW	respiratory disorder; se.				
XX					
OS	Homo sapiens.				
XX					
PN	WO200162927-A2.				
PD	30-AUG-2001.				
XX					
PF	21-FEB-2001; 2001WO-US006059.				
XX					
PR	24-FEB-2000; 2000US-0184693P.				
PR	24-FEB-2000; 2000US-0184697P.				
PR	24-FEB-2000; 2000US-0184698P.				
PR	24-FEB-2000; 2000US-0184768P.				
PR	24-FEB-2000; 2000US-0184769P.				
PR	24-FEB-2000; 2000US-0184770P.				
PR	24-FEB-2000; 2000US-0184771P.				
PR	24-FEB-2000; 2000US-0184772P.				
PR	24-FEB-2000; 2000US-0184773P.				
PR	24-FEB-2000; 2000US-0184774P.				
PR	24-FEB-2000; 2000US-0184776P.				
PR	24-FEB-2000; 2000US-0184797P.				
PR	24-FEB-2000; 2000US-0184797P.				
PR	24-FEB-2000; 2000US-0184813P.				
PR	24-FEB-2000; 2000US-0184837P.				
PR	24-FEB-2000; 2000US-0184841P.				
PR	24-FEB-2000; 2000US-0185213P.				
PR	24-FEB-2000; 2000US-0185216P.				
PR	12-MAY-2000; 2000US-0203785P.				
PR	13-MAY-2000; 2000US-0204226P.				
PR	16-MAY-2000; 2000US-0204525P.				
PR	16-MAY-2000; 2000US-0204821P.				
PR	16-MAY-2000; 2000US-0204908P.				
PR	16-MAY-2000; 2000US-0205232P.				
PR	17-MAY-2000; 2000US-0204815P.				
PR	17-MAY-2000; 2000US-0204863P.				

PR 17-MAY-2000; 2000US-0205221P.  
 PR 17-MAY-2000; 2000US-0205285P.  
 PR 17-MAY-2000; 2000US-0205286P.  
 PR 17-MAY-2000; 2000US-0205287P.  
 PR 17-MAY-2000; 2000US-0205323P.  
 PR 17-MAY-2000; 2000US-0205324P.  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Panzer SR, Spino PA, Banville SC, Shah P, Chalup MS, Chang SC,  
 PI Chen A, D'ea SA, Amesley S, Dahl CR, Dam TC, Daniels SE, Dufour GE,  
 PI Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL, Liu TY,  
 PI Roseberry AM, Rosen BH, Russo PD, Stockbrecher TK, Daiffo A,  
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W,  
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
 XX WPI, 2001-502867/55.  
 DR P-PSDB; AAU19557.  
 PT Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.  
 XX  
 PS Claim 1; Page 363-364; 522pp; English.  
 XX  
 CC The invention relates to polynucleotides (I) encoding diagnostic and  
 CC therapeutic (DTHP) polypeptides (II), which include e.g. enzymes, and  
 CC proteins involved in growth and development and receptors. (I) and (II)  
 CC may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate DTHP expression. For example, (I) and (II)  
 CC may be used to treat disorders associated with decreased polypeptide  
 CC expression by rectifying mutations or deletions in a patient's genome,  
 CC that affect the activity of the DTHPs, by expressing inactive proteins  
 CC or supplementing the patient's own production of them. (I) and (II) may  
 CC be used to treat diseases, for example, cell proliferative disorder,  
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
 CC (I) may be used to produce the DTHPs, by inserting the nucleic acids  
 CC into a host cell and culturing the cell to express the protein. (I) and  
 CC its complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids in  
 CC samples, and therefore which patients may be in need of restorative  
 CC therapy. (II) may also be used as antigens in the production of  
 CC antibodies against DTHPs and in assays to identify modulators of DTHP  
 CC expression and activity. The anti-DTHP antibodies and antagonists may  
 CC also be used to down regulate expression and activity. The anti-DTHP  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of DTHPs in samples (e.g. by enzyme linked immunosorbent assay  
 CC (ELISA)). AAS30986-AAS31156 represent human diagnostic and therapeutic  
 CC (DTHP) polynucleotides of the invention  
 XX  
 SQ Sequence 1171 BP; 237 A; 375 C; 356 G; 201 T; 0 U; 2 Other;  
 Query Match 99.4%; Score 1154.2; DB 4; Length 1171;  
 Best Local Similarity 99.6%; Pred. No. 2.8e-261;  
 Matches 1156; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 GGAATATGATGCTGGGAGCGAGCTGCTGCGGCTGCGGAGGAAAGCGGAGAC 60  
 DB 8 GGAATATGATGCTGGGAGCGAGCTGCTGCGGCTGCGGAGGAAAGCGGAGAC 67  
 QY 61 GCGAGAGCGCGCGAGAGCTTTGCTCGGAGCGCCCTGAGAGTGGCGGAGCGCG 120  
 DB 68 GCGAGAGCGCGCGAGAGCTTTGCTCGGAGCGCCCTGAGAGTGGCGGAGCGCG 127  
 QY 121 AGGAGTAAACCAATGATCCCTGGGTGCTCTGGCTGTGCTCCCTGTGCTGTGAC 180  
 DB 128 AGGAGTAAACCAATGATCCCTGGGTGCTCTGGCTGTGCTCCCTGTGCTGTGAC 187  
 QY 181 CCACTGCTTGGCGCTTGTGCTGCGAGGAGACTTCGGAAGGCTCCCTCAATGATGCTGC 240  
 DB 188 CCACTGCTTGGCGCTTGTGCTGCGAGGAGACTTCGGAAGGCTCCCTCAATGATGCTGC 247  
 QY 241 AGCTGCTGCGCCCGAGGCGCAACCGGCGCCCGCAGAGAGCCGAGGCGCTCAGAAATG 300

DB 248 AGCTGCTGCGCCCGAGGCGCAACCGGCGCCCGCAGAGAGCCCGAGGCGCTCAGAAATG 307  
 QY 301 ATGGAGAGATGAGGCTTTCTGCGAAGAGAGGCGCAAGATGAGACAGCGCGAG 360  
 DB 308 ATGGAGAGATGAGGCTTTCTGCGAAGAGAGGCGCAAGATGAGACAGCGCGAG 367  
 QY 361 GACAGCGAGAGAGAGGCTTCCAGCTGGCGGAGCAGATTAACCGGAGAAAGCCAGAGCCAAAG 420  
 DB 368 GACAGCGAGAGAGAGAGGCTTCCAGCTGGCGGAGCAGATTAACCGGAGAAAGCCAGAGCCAAAG 427  
 QY 421 GCGAAGCGGCGCGCATTTGGCGGCGGCTGGCGCCCGCTGCGCCCAAGGGGTCAACGGTACC 480  
 DB 428 GCGAAGCGGCGCGCATTTGGCGGCGGCTGGCGCCCGCTGCGCCCAAGGGGTCAACGGTACC 487  
 QY 481 CCGGAGAGAGATGAGCAACCGAGGAGAGAGGCGCCCAAGGGGCAAGAAAGGCGCGAGG 540  
 DB 488 CCGGAGAGAGATGAGCAACCGAGGAGAGAGGCGCCCAAGGGGCAAGAAAGGCGCGAGG 547  
 QY 541 CTCGAGAGCGCTGAGCTGGCAGTGGCAGTGGCCATACCAAGTCAAGCTTCTCGTGGCAGTG 600  
 DB 548 CTCGAGAGCGCTGAGCTGGCAGTGGCAGTGGCCATACCAAGTCAAGCTTCTCGTGGCAGTG 607  
 QY 601 ACCAAGAGCTTACCAGCGGAGCGGCTGCCATCAAGTTTGAACAATTTGATGAACGAG 660  
 DB 608 ACCAAGAGCTTACCAGCGGAGCGGCTGCCATCAAGTTTGAACAATTTGATGAACGAG 667  
 QY 661 GGTGGCCACATCAATGTTTCCAGCGGCAAGTTGTCGGCGGCTGGCGGATCTACTAC 720  
 DB 668 GGTGGCCACATCAATGTTTCCAGCGGCAAGTTGTCGGCGGCTGGCGGATCTACTAC 727  
 QY 721 TTCACTTACGATCACTACGCTGGCGCAACAGCACTGGCCATCGGCTGTGCAACAGG 780  
 DB 728 TTCACTTACGATCACTACGCTGGCGCAACAGCACTGGCCATCGGCTGTGCAACAGG 787  
 QY 781 CAGTACCGCATCGGAGCTTTGATGCCACACAGCGGCAACGATGAGGCTCAGGCTCC 840  
 DB 788 CAGTACCGCATCGGAGCTTTGATGCCACACAGCGGCAACGATGAGGCTCAGGCTCC 847  
 QY 841 ACCATCTGGCTGCAAGAGGAGTGAAGAGTTTGGTCGAGATCTTCACTCAGAGCAG 900  
 DB 848 ACCATCTGGCTGCAAGAGGAGTGAAGAGTTTGGTCGAGATCTTCACTCAGAGCAG 907  
 QY 901 AACGGGCTCTTATAGACCTTTACTGAGACAGAGCTTTTACGGGCTTCTAATCTAT 960  
 DB 908 AACGGGCTCTTATAGACCTTTACTGAGACAGAGCTTTTACGGGCTTCTAATCTAT 967  
 QY 961 GCCGACCAAGATGACCCCAACGAGATGACATGCCACGCGGCTCTCAGGCAAGGAA 1020  
 DB 968 GCCGACCAAGATGACCCCAACGAGATGACATGCCACGCGGCTCTCAGGCAAGGAA 1027  
 QY 1021 CAGGCTTGGAGCTTGGGCTTACAGAGCAAGAGCCCAACATGAGGCTGGGGTGGGG 1080  
 DB 1028 CAGGCTTGGAGCTTGGGCTTACAGAGCAAGAGCCCAACATGAGGCTGGGGTGGGG 1087  
 QY 1081 GTGAGTGAAGCGGTTTACGCTCAGGCTCAAGCTCTCGGCTTTTTCCTCCCTCAT 1140  
 DB 1088 GTGAGTGAAGCGGTTTACGCTCAGGCTCAAGCTCTCGGCTTTTTCCTCCCTCAT 1147  
 QY 1141 AAATCCAACTTTTATTC 1161  
 DB 1148 AAATCCAACTTTTATTC 1168  
 RESULT 6  
 ACD03627  
 ID ACD03627 standard; cDNA; 1178 BP.  
 XX ACD03627;  
 AC ACD03627;  
 XX  
 DT 01-AUG-2003 (first entry)  
 XX Novel human GPCR related protein NOV5a cDNA.

XX Human; G-protein coupled receptor related protein; GPCR related protein;  
 KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;  
 KW immunomodulator; anti-HIV; anorectic; antisthmatic; haemostatic;  
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;  
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;  
 KW diabetes; immune disorder; AIDS; obesity; asthma;  
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;  
 KW infection; multiple sclerosis; cancer-associated cachexia;  
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;  
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;  
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 PN WO200299116-A2.  
 PD 12-DEC-2002.  
 XX  
 PF 04-JUN-2002; 2002WO-US017428.  
 XX  
 PR 04-JUN-2001; 2001US-0295607P.  
 PR 04-JUN-2001; 2001US-0295661P.  
 PR 06-JUN-2001; 2001US-0296404P.  
 PR 06-JUN-2001; 2001US-0296418P.  
 PR 14-JUN-2001; 2001US-0298285P.  
 PR 15-JUN-2001; 2001US-0298556P.  
 PR 21-JUN-2001; 2001US-0299949P.  
 PR 26-JUN-2001; 2001US-0300883P.  
 PR 26-JUN-2001; 2001US-0301550P.  
 PR 13-AUG-2001; 2001US-0311972P.  
 PR 27-AUG-2001; 2001US-0315071P.  
 PR 29-AUG-2001; 2001US-0315660P.  
 PR 14-SEP-2001; 2001US-0322293P.  
 PR 17-SEP-2001; 2001US-0322706P.  
 PR 14-DEC-2001; 2001US-0341186P.  
 PR 28-FEB-2002; 2002US-0361189P.  
 PR 12-MAR-2002; 2002US-0363673P.  
 PR 12-MAR-2002; 2002US-0363676P.  
 PR 03-JUN-2002; 2002US-00363676.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Anderson DW, Baumgartner JC, Boldog FL, Caeman SJ, Edinger SR, Li L;  
 PI Gangoli EA, Gerlach VL, Gorman L, Guo X, Hjalte T, Kekuda R, Li L;  
 PI McDonough JR, Malyanar UM, Miller I, Padigaru M, Paturajan M;  
 PI Pena CE, Raschelli L, Shinkets RA, Stone DJ, Spytek KA, Vernet CM;  
 PI Voss EZ, Zehusen BD;  
 XX  
 DR WPI; 2003-140627/13.  
 DR P-PSDB; ABU99123.  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for preventing or  
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 20; Page 108-109; 332pp; English.  
 XX  
 CC The invention describes an isolated polypeptide (1) comprising any of 27  
 CC 118-961 residue amino acid sequences, given in the specification, a  
 CC mature form of them, a sequence that is at least 95 % identical to them,  
 CC or a sequence having one or more conservative substitutions in them. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease selected from a pathology  
 CC associated with the polypeptide. The NOVX polypeptides, polynucleotides  
 CC and antibodies are useful in treating or preventing NOVX-associated  
 CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune  
 CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's  
 CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-  
 CC associated cachexia, and other wasting disorders associated with chronic  
 CC diseases. The nucleic acids and polypeptides may also be used as targets

CC for the identification of small molecules that modulate or inhibit e.g.  
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,  
 CC wound healing and angiogenesis, in gene therapy, in generation of  
 CC antibodies that bind immunospecifically to NOVX substances for use in  
 CC therapeutic or diagnostic methods. The nucleic acids are further used as  
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive  
 CC medicine, and pharmacogenomics. The polypeptides are also useful as  
 CC vaccines. This sequence encodes a novel human G-protein coupled receptor  
 CC related protein NOV  
 XX  
 XX Sequence 1178 BP; 244 A; 377 C; 355 G; 202 T; 0 U; 0 Other;  
 XX  
 XX Query Match 98.9%; Score 1148.4; DB 8; Length 1178;  
 XX Best Local Similarity 99.8%; Pred. No. 6.4e-260;  
 XX Matches 1160; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 GGAATACTATGCTGGGGCCGACGCTGCGCCGCTGCTGCGTGAAGAAAGCCGGAC 60  
 Db 7 GGAATACTATGCTGGGGCCGACGCTGCGCCGCTGCTGCGTGAAGAAAGCCGGAC 66  
 QY 61 GCGGAGCCCGCCGAGAGCTTCTTGTCTCGGAGCCGCCCTGAGCTGGCGGACGCCG 120  
 Db 67 GCGGAGCCCGCCGAGAGCTTCTTGTCTCGGAGCCGCCCTGAGCTGGCGGAGCCGCG 126  
 QY 121 AGGATAACACCATGATCCCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 Db 127 AGGATAACACCATGATCCCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186  
 QY 181 CCACTGCTTGGCGCTTGTGCTGCGAGGAGCTTCGGAAGGCTCCCTCACTGTGTGC 240  
 Db 187 CCACTGCTTGGCGCTTGTGCTGCGAGGAGCTTCGGAAGGCTCCCTCACTGTGTGC 246  
 QY 241 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 Db 247 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306  
 QY 301 ATGGGAGGAATGGGCTTCCGGAAGAGCGGCAAGATGAGACAGAGCGGAGCGGAGG 360  
 Db 307 ATGGGAGGAATGGGCTTCCGGAAGAGCGGCAAGATGAGACAGAGCGGAGCGGAGG 366  
 QY 361 GACAGCGGAGGAGGAGGCTCACTGCGGAGCAAGTAAACCGGGAAGGACAGACCAAG 420  
 Db 367 GACAGCGGAGGAGGAGGCTCACTGCGGAGCAAGTAAACCGGGAAGGACAGACCAAG 426  
 QY 421 GGCAGAGCCGCGGCTTGGGCGGCGCTGCGCCCGCTGCGCCCGCTGCGGCTCAAGCTA 480  
 Db 427 GGCAGAGCCGCGGCTTGGGCGGCGCTGCGCCCGCTGCGCCCGCTGCGGCTCAAGCTA 486  
 QY 481 CCGGGAAGGATGGCAACAGGCAAGAGGCGCCCAAGGCAAGAGGGAAGCCAGGC 540  
 Db 487 CCGGGAAGGATGGCAACAGGCAAGAGGCGCCCAAGGCAAGAGGGAAGCCAGGC 546  
 QY 541 CTCGAGGCGCTGCACTGCTGCGAGTGGCAATCAAGTCAAGTCTGCGTGGCAGTG 600  
 Db 547 CTCGAGGCGCTGCACTGCTGCGAGTGGCAATCAAGTCAAGTCTGCGTGGCAGTG 606  
 QY 601 ACCAAGAGTACCAACGAGGAGCGGCTGCCATCAAGTTTGAACAAGATTGTGATGAC 660  
 Db 607 ACCAAGAGTACCAACGAGGAGCGGCTGCCATCAAGTTTGAACAAGATTGTGATGAC 666  
 QY 661 GGTGGCACTACATGCTTTCAGTGGCAATGCTGCGGCGGCTGCGGATTTACTAC 720  
 Db 667 GGTGGCACTACATGCTTTCAGTGGCAATGCTGCGGCGGCTGCGGATTTACTAC 726  
 QY 721 TTCACTACGACATCACTGCTGCGCAACAGCACTGCGCTGCTGCTGCTGCTGCTGCT 780  
 Db 727 TTCACTACGACATCACTGCTGCGCAACAGCACTGCGCTGCTGCTGCTGCTGCTGCT 786  
 QY 781 CAGTACCGATCGGAGCTTTGATGCAACAGCGCAACGATGCTGCTGCTGCTGCTGCT 840  
 Db 787 CAGTACCGATCGGAGCTTTGATGCAACAGCGCAACGATGCTGCTGCTGCTGCTGCTGCT 846  
 QY 841 ACATCTGCTGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900

Db 847 ACCATCTGCTCTCAAGAGGGGAGAGAGTTGGCTGAGAGCTTCTACTAGAGAG 906  
 Qy 901 AACGGGCTCTTATAGACCTTACTGAGACAGACCTCTTTACGGGCTTCTATCTAT 960  
 Db 907 AACGGGCTCTTATAGACCTTACTGAGACAGACCTCTTTACGGGCTTCTATCTAT 966  
 Qy 961 GCGGACGAGATGACCCCAAGAGATATAGACAGGCGGCTCTCCAGAGGAGAA 1020  
 Db 967 GCGGACGAGATGACCCCAAGAGATATAGACAGGCGGCTCTCCAGAGGAGAA 1026  
 Qy 1021 CAAGCTTCTGACTTGGGCTTACAGACAGACCCCAACTGTAGGCTGGGCTGGGAG 1080  
 Db 1027 CAAGCTTCTGACTTGGGCTTACAGACAGACCCCAACTGTAGGCTGGGCTGGGAG 1086  
 Qy 1081 GTGAGTGAAGGGGTTCTAGGCTCAGGCTCAGCTCTCTCGGCTC-TTTTTTCCCTTCAT 1139  
 Db 1087 GTGAGTGAAGGGGTTCTAGGCTCAGGCTCAGCTCTCTCGGCTC-TTTTTTCCCTTCAT 1146  
 Qy 1140 TAAATCCAACCTTTTATCA 1161  
 Db 1147 TAAATCCAACCTTTTATCA 1168  
 RESULT 7  
 ID ADM93810 standard; DNA, 1178 BP.  
 AC ADM93810;  
 XX 17-JUN-2004 (first entry)  
 DE DNA encoding human NOV protein #15.  
 XX  
 KM gene therapy; vaccine; ds; gene; NOX; cancer;  
 KM neurodegenerative disorder; Parkinson's disease; metabolic disorder;  
 KM diabetes; obesity; immune related disorder; tissue typing; human.  
 OS Homo sapiens.  
 XX  
 PN US2004009480-A1.  
 PD 15-JAN-2004.  
 XX  
 PF 03-JUN-2002; 2002US-00162335.  
 XX  
 PR 04-JUN-2001; 2001US-0295607P.  
 PR 04-JUN-2001; 2001US-029561P.  
 PR 06-JUN-2001; 2001US-0296404P.  
 PR 06-JUN-2001; 2001US-0296418P.  
 PR 11-JUN-2001; 2001US-0297414P.  
 PR 12-JUN-2001; 2001US-0297567P.  
 PR 14-JUN-2001; 2001US-0298285P.  
 PR 15-JUN-2001; 2001US-0299949P.  
 PR 21-JUN-2001; 2001US-0300883P.  
 PR 26-JUN-2001; 2001US-0300883P.  
 PR 28-JUN-2001; 2001US-0301550P.  
 PR 13-AUG-2001; 2001US-0311972P.  
 PR 27-AUG-2001; 2001US-0315069P.  
 PR 29-AUG-2001; 2001US-0315071P.  
 PR 14-SEP-2001; 2001US-0322939P.  
 PR 17-SEP-2001; 2001US-0322706P.  
 PR 14-DEC-2001; 2001US-0341186P.  
 PR 28-FEB-2002; 2002US-0363673P.  
 PR 12-MAR-2002; 2002US-0363673P.  
 PR 12-MAR-2002; 2002US-0363676P.  
 XX  
 PA (ANDE/) ANDERSON D W.  
 PA (BAUM/) BAUMGARTNER J C.  
 PA (BOLD/) BOLDOG F L.  
 PA (CASW/) CASMAN S J.  
 PA (EDIN/) EDINGER S R.

PA (GANG/) GANGOLLI E A.  
 PA (GERL/) GERLACH V.  
 PA (GORM/) GORMAN L.  
 PA (GUOX/) GUO X S.  
 PA (HJAL/) HJALT T.  
 PA (KEKU/) KEKUDA R.  
 PA (LIL/) LI L.  
 PA (MACD/) MACDOUGALL J R.  
 PA (MALY/) MALYANKAR U M.  
 PA (MILT/) MILLET I.  
 PA (PADI/) PADIGARU M.  
 PA (PATU/) PATTURAJAN M.  
 PA (PEN/) PENNA C E A.  
 PA (RAST/) RASTELLI L.  
 PA (SHIM/) SHIMKETS R A.  
 PA (STON/) STONE D J.  
 PA (SPYT/) SPYTEK K A.  
 PA (VERN/) VERNET C A M.  
 PA (VOSS/) VOSS E Z.  
 PA (ZERH/) ZERHUSEN B D.  
 PI Anderson DM, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR, Gangolli EA, Gerlach V, Gorman L, Guo XS, Hjal T, Kekuda R, Li L, MacDougall JR, Malyankar UM, Millet I, Padigar M, Patturajan M, Pena CE, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CM, Voss EZ, Zerhusen BD;  
 PI WPI, 2004-090456/09.  
 DR P-PDB; ADM93811.  
 XX  
 PT New NOVX polypeptide, useful for preparing a composition for treating or preventing e.g., cancer, neurodegenerative disorders such as Parkinson's disease, or metabolic disorders such as diabetes or obesity, or for tissue typing.  
 PT  
 PT  
 PT  
 PT  
 PS Disclosure; SEQ ID NO 29; 202pp; English.  
 XX  
 XX  
 CC The invention relates to an isolated NOVX polypeptide. The polypeptide is useful for preparing a composition for treating or preventing a pathology associated with NOVX polypeptide e.g. cancer, neurodegenerative disorders such as Parkinson's disease, metabolic disorders such as diabetes or obesity or immune related disorders or for tissue typing. The present sequence represents DNA encoding a human NOV protein.  
 CC  
 CC  
 SQ Sequence 1178 BP; 244 A; 382 C; 350 G; 202 T; 0 U; 0 Other;  
 Query Match 98.2%; Score 1140.4; DB 12; Length 1178;  
 Best Local Similarity 99.4%; Pred. No. 4.9e-258;  
 Matches 1155; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
 Qy 1 GGAATACTATGCTGGGGGCGAGGCTGCGGGCTGCGGCTGAGGAAAGCGGGAG 60  
 Db 7 GGAATACTATGCTGGGGGCGAGGCTGCGGGCTGCGGCTGAGGAAAGCGGGAG 66  
 Qy 61 GCGAGGCCCGCGAGAGCTTCTTGGCTCCGAGCGCCCTGAGAGTGGCGGAGCCGCG 120  
 Db 67 GCGAGGCCCGCGAGAGCTTCTTGGCTCCGAGCGCCCTGAGAGTGGCGGAGCCGCG 126  
 Qy 121 AGGGTAACAACATGATCCCTGAGGCTCTGAGCTGAGCTGAGCTGAGCTGAG 180  
 Db 127 AGGGTAACAACATGATCCCTGAGGCTCTGAGCTGAGCTGAGCTGAGCTGAG 186  
 Qy 181 CCACTGCTGGGCGCTTGTCTCCAGAGATTCCGAAAGGCTCCCTCAACTGCTGTC 240  
 Db 187 CCACTGCTGGGCGCTTGTCTCCAGAGATTCCGAAAGGCTCCCTCAACTGCTGTC 246  
 Qy 241 AGCTGCTGAGCGCCCAAGGCGCCAGCGCCCGCCAGAGCGCCCTGAGGATG 300  
 Db 247 AGCTGCTGAGCGCCCAAGGCGCCAGCGCCCGCCAGAGCGCCCTGAGGATG 306  
 Qy 301 ATGGAGCAATGGGCTTCTTGGCAAGAGCGGCAAGATGAGACAGAGCGGCGAG 360  
 Db 307 ATGGAGCAATGGGCTTCTTGGCAAGAGCGGCAAGATGAGACAGAGCGGCGAG 366



```

Qy 361 GACAGCGGAGAGAAAGTCACTGGCCGGAACGCTAACCGGGAAAGCCAGACCAAG 420
Db 367 GACAGCGGAGAGAAAGTCACTGGCCGGAACGCTAACCGGGAAAGCCAGACCAAG 426
Qy 421 GGCAGAGCCGGGGCCATTGGGCGGGCTGGCCCGGCTGGCCCAAGAGGGGTCAACGGTACC 480
Db 427 GGCAGAGCCGGGGCCATTGGGCGGGCTGGCCCGGCTGGCCCAAGAGGGGTCAACGGTACC 486
Qy 481 CCCGGAGAGCATGGGCACACAGGCAAGAGAGGGGCCAAGGCGAAGAGAGGAGCCAGGC 540
Db 487 CCCGGAGAGCATGGGCACACAGGCAAGAGAGGGGCCAAGGCGAAGAGAGGAGCCAGGC 546
Qy 541 CTCGCAAGGCGCCCTGAGCTGTGGCACTGGCCATCAAGTCAAGTCTTCCGCTGGCAATG 600
Db 547 CTCGCAAGGCGCCCTGAGCTGTGGCACTGGCCATCAAGTCAAGTCTTCCGCTGGCAATG 606
Qy 601 ACCAAGAGCTACCAACGAGGAGCGGCTGCCATCAAGTCTTGAACAAGATTCTGATGAACGAG 660
Db 607 ACCAAGAGCTACCAACGAGGAGCGGCTGCCATCAAGTCTTGAACAAGATTCTGATGAACGAG 666
Qy 661 GGTGGCCACTAACAATGCTTCCAGCGGCAAGTTCTGCGGCGTGCCTGGAGATCTACTAC 720
Db 667 GGTGGCCACTAACAATGCTTCCAGCGGCAAGTTCTGCGGCGTGCCTGGAGATCTACTAC 726
Qy 721 TTCACTTAACGACATCAAGCTGGCCAAAGACCTGGGCAATCGGCTGTGTGCAACAACGCG 780
Db 727 TTCACTTAACGACATCAAGCTGGCCAAAGACCTGGGCAATCGGCTGTGTGCAACAACGCG 786
Qy 781 CAGTACCGCATCCGGAACCTTGTATGATGCCAACACCGGCAACAGATGTGGCTCAGAGCTCC 840
Db 787 CAGTACCGCATCCGGAACCTTGTATGATGCCAACACCGGCAACAGATGTGGCTCAGAGCTCC 846
Qy 841 ACCATCTGGCTCTCAAGCAGAGGCTGACGAGATTTGGCTGCAAGATCTTCTACTCAGAGAG 900
Db 847 ACCATCTGGCTCTCAAGCAGAGGCTGACGAGATTTGGCTGCAAGATCTTCTACTCAGAGAG 906
Qy 901 AACGGGCTCTTCTATGACCTTACTGACACAGACCTCTTAAAGGCTTCTTAATCAT 960
Db 907 AACGGGCTCTTCTATGACCTTACTGACACAGACCTCTTAAAGGCTTCTTAATCAT 966
Qy 961 GCCAGCAGATGAGACCCCAAGAGATATGACATGCCAGCGGCTCTCAAGCAGAGGAA 1020
Db 967 GCCAGCAGATGAGACCCCAAGAGATATGACATGCCAGCGGCTCTCAAGCAGAGGAA 1026
Qy 1021 CAACTTCTGAGCTTGGCTTACAGAGCAAGACCCCAACATGTAGGCTGGGGGTGGGGG 1080
Db 1027 CAACTTCTGAGCTTGGCTTACAGAGCAAGACCCCAACATGTAGGCTGGGGGTGGGGG 1086
Qy 1081 GTGAGTGAAGGGTCTTAGCTTACGCTCAAGCTCTCTCGGCTC-TTTTTTCCCTTCAAT 1139
Db 1087 GTGAGTGAAGGGTCTTAGCTTACGCTCAAGCTCTCTCGGCTCTTTTTTCCCTTCAAT 1146
Qy 1140 TAAATCCAAACCTTTTATTCA 1161
Db 1147 TAAATCCAAACCTTTTATTCA 1168

```

RESULT 8  
AB054564  
ID AB054564 standard; cDNA; 1171 BP.

AC AB054564,  
DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HDHB13 cDNA, SEQ ID NO:444.  
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;

KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive; gene; ss.  
OS Homo sapiens.  
EN WO200200677-A1.  
XX 03-JAN-2002.  
PD 07-JUN-2001; 2001WO-US018569.  
XX 07-JUN-2000; 2000US-0209467P.  
PR 07-JUN-2000; 2000US-0209467P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Birse CE, Rosen CA;  
PI WPI, 2002-147878/19.  
DR P-PSDB; ABP41487.  
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
PT cancer), immune disorders, cardiovascular disorders and neurological  
PT diseases.  
PS Claim 1; SEQ ID NO 444; 2922pp; English.  
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
XX ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired  
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
CC respiratory disorders, neurological disorders, gastrointestinal disorders  
CC and urinary system disorders. Ovarian antigen polypeptides and  
CC polynucleotides may also be used in screening for compounds which  
CC modulate ovarian antigen expression or activity. The polynucleotides may  
CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents cDNA encoding a human ovarian antigen of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 1171 BP; 251 A; 374 C; 347 G; 199 T; 0 U; 0 Other;

Query Match 96.5%; Score 1120.8; DB 6; Length 1171;  
Best Local Similarity 99.2%; Pred. No. 2e-253;  
Matches 1137; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

```

Qy 18 GCCAGCGCTGCGCCGGCTGCTGCGGAGAAAGCCGGAGCGCGAGCC--CGCGGA 75
Db 1 GCCAGCGCTGCGCCGGCTGCTGCGGAGAAAGCCGGAGCGCGAGCCCGGACAGAG 60
Qy 76 GAGCTTCTTGTCTCCGAGCGCCCTGAGCGTGGCGGACGCGGAGGTTAACCAATG 135
Db 61 GAGCTTCTTGTCTCCGAGCGCCCTGAGCGTGGCGGACGCGGAGGTTAACCAATG 120

```

QY 136 ATCCCTGAGTCTCTTGCCCTGCTCCCTGCTGCTGACCCAGTCTTGAGCC 195  
 DB 121 ATCCCTGAGTCTCTTGCCCTGCTCCCTGCTGCTGACCCAGTCTTGAGCC 180  
 QY 196 TTGCTGAGGAACTTCGGAAGAGCTCCCTCACTGCTGCTGAGCTGCTGCTG 255  
 DB 181 TTGCTGAGGAACTTCGGAAGAGCTCCCTCACTGCTGCTGAGCTGCTGCTG 240  
 QY 256 CAGGAGCCACCCGAGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGGAGCC 315  
 DB 241 CAGGAGCCACCCGAGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGGAGCC 300  
 QY 316 TTGCTGAGGAACTTCGGAAGAGCTCCCTCACTGCTGCTGAGCTGCTGCTG 375  
 DB 301 TTGCTGAGGAACTTCGGAAGAGCTCCCTCACTGCTGCTGAGCTGCTGCTG 360  
 QY 376 GGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 435  
 DB 361 GGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
 QY 436 ATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 495  
 DB 421 ATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
 QY 496 ACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 555  
 DB 481 ACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 QY 556 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 615  
 DB 541 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 QY 616 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 675  
 DB 601 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
 QY 676 GCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 735  
 DB 661 GCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
 QY 736 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 795  
 DB 721 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
 QY 796 ACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 855  
 DB 781 ACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 QY 856 AAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915  
 DB 841 AAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
 QY 916 GACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975  
 DB 901 GACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 QY 976 CCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035  
 DB 961 CCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
 QY 1036 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1095  
 DB 1021 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 QY 1096 CTAGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1155  
 DB 1081 CTAGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
 QY 1156 TATTCGA 1161  
 DB 1141 TATTCGA 1146

RESULT 9  
 ADF94944/C  
 ID ADF94944 standard; cDNA; 1211 BP.  
 XX  
 AC ADF94944;  
 XX  
 DT 26-FEB-2004 (first entry)  
 XX  
 DE Human gene 11-derived ACRP30-like cDNA, SEQ ID NO:40.  
 XX  
 XX Human; ACRP30-like; adipocyte complement related protein 30-like;  
 KW complement factor C1q homologue; cerebellin homologue;  
 KW hibernation-specific protein homologue;  
 KW tumour necrosis factor-alpha homologue; TNF-alpha homologue;  
 KW drug screening; diabetes mellitus; type 2 diabetes; insulin resistance;  
 KW obesity; metabolic disorder; endocrine disorder; antidiabetic; anorectic;  
 KW gene therapy; chromosome 11q11; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN M02003031586-A2.  
 XX  
 PD 17-APR-2003.  
 XX  
 PF 11-OCT-2002; 2002MO-US032432.  
 XX  
 PR 12-OCT-2001; 2001US-0328419P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (BLON/) BLONDEL O.  
 PA (RUB/) RUBEN S M.  
 XX  
 PI Blondel O, Ruben SM;  
 XX  
 DR WPJ, 2003-381707/36.  
 DR P-PSDB; ADF94885.  
 XX  
 PT New ACRP30-like nucleic acid molecules and polypeptides, useful for  
 PT preventing, treating or ameliorating a medical condition, such as insulin  
 PT resistance or diabetes mellitus.  
 XX  
 PS Claim 1; SEQ ID NO 40; 602pp; English.  
 XX  
 CC The invention relates to novel human adipocyte complement related protein  
 CC 30 (ACRP30)-like cDNAs derived from 20 genes, and their encoded proteins.  
 CC The secretion of ACRP30 is promoted by insulin, and ACRP30 acts in a  
 CC similar manner to insulin, such as reducing blood glucose levels and  
 CC reducing levels of plasma free fatty acids. ACRP30 is homologous to  
 CC complement factor C1q, cerebellin, and hibernation-specific proteins HP-  
 CC 20, HP-25, and HP-27, and has a C-terminal globular domain similar in  
 CC crystal structure to C1q and tumour necrosis factor-alpha (TNF-alpha),  
 CC the latter of which has been implicated in insulin resistance in obesity  
 CC and type 2 diabetes. Additionally ACRP30 is highly expressed in  
 CC adipocytes during adipogenesis, and may also have immune-related  
 CC functions. ACRP30-like proteins may therefore be useful in the treatment  
 CC of insulin-related disorders such as diabetes and obesity. The invention  
 CC also relates to recombinant vectors and host cells comprising an ACRP30-  
 CC like nucleic acid; the recombinant production of ACRP30-like proteins;  
 CC antibodies specific for ACRP30-like proteins; use of ACRP30-like nucleic  
 CC acids or proteins in preventing, treating or ameliorating a medical  
 CC condition in an individual; methods of diagnosing a pathological  
 CC condition, or a susceptibility to a pathological condition; methods of  
 CC screening for binding partners of ACRP30-like proteins; and methods of  
 CC screening for modulators of ACRP30-like activity. The ACRP30-like nucleic  
 CC acids and proteins, and methods of the invention are useful for  
 CC preventing, treating or ameliorating a medical condition such as diabetes  
 CC mellitus, insulin resistance, obesity, and other metabolic and endocrine  
 CC disorders. The present sequence is related to the invention.  
 XX  
 SQ Sequence 1211 BP; 209 A; 356 C; 380 G; 264 T; 0 U; 2 Other;

Query Match 93.4%; Score 1084; DB 10; Length 1211;

Best Local Similarity 99.8%; Pred. No. 8.8e-245;  
Matches 1084; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy 76 GAGCTTCTTGTCTCGGACGCCCCCTGACGTGCGGGGACGCGGAGGTAAACCAATG 135
Db 1111 GAGCTTCTTGTCTCGGACGCCCCCTGACGTGCGGGGACGCGGAGGTAAACCAATG 1052

Qy 136 ATCCCCGGGGTCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 195
Db 1051 ATCCCCGGGGTCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 992

Qy 196 TTTGCTGCGAGGACCTTCGGGAAAGGCTCCCTCAACTGAGCTGAGCTGCGCTGCGCTG 255
Db 991 TTTGCTGCGAGGACCTTCGGGAAAGGCTCCCTCAACTGAGCTGAGCTGCGCTGCGCTG 932

Qy 256 CAGGGCCACCCGCCCCCGGAGGCCCTTCAAGATGATGGACGAATGGG 315
Db 931 CAGGGCCACCCGCCCCCGGAGGCCCTTCAAGATGATGGAGCAATGGG 872

Qy 316 TTTCTTGGCAAAAGCGGCAAGATGACACGCGGACCGGGGGGACAGGGAGAGAA 375
Db 871 TTTCTTGGCAAAAGCGGCAAGATGACACGCGGACCGGGGGGACAGGGAGAGAA 812

Qy 376 GGTCCACTGGCGGAGAGGTAAACGGGAAAGCCAGGACCAAGGGGAAAGCCGGGGCC 435
Db 811 GGTCCACTGGCGGAGAGGTAAACGGGAAAGCCAGGACCAAGGGGAAAGCCGGGGCC 752

Qy 436 ATTGGGCGGGCTGGCCCCCGTGGCCCCCAAGGGGATCAAGGTACCCCGGGAGCATGGC 495
Db 751 ATTGGGCGGGCTGGCCCCCGTGGCCCCCAAGGGGATCAAGGTACCCCGGGAGCATGGC 692

Qy 496 ACACCAAGCAAGAAAGGGGCCCCAAGGGCAAAAGGGAGCCAGGCTTCCAGGGCCCTGCG 555
Db 691 ACACCAAGCAAGAAAGGGGCCCCAAGGGCAAAAGGGAGCCAGGCTTCCAGGGCCCTGCG 632

Qy 556 AGCTGGGAGTGGGCAATACCAATGCTTCTCGGTGGAGTGAACCAAGAGCTACCCA 615
Db 631 AGCTGGGAGTGGGCAATACCAATGCTTCTCGGTGGAGTGAACCAAGAGCTACCCA 572

Qy 616 CGGGAGCGGCTGCCCATCAAGTTTGACAAATTTGATGAACGAAGGTGGCCACTACAT 675
Db 571 CGGGAGCGGCTGCCCATCAAGTTTGACAAATTTGATGAACGAAGGTGGCCACTACAT 512

Qy 676 GCTTCCAGCGGCAAGTTGCTGCGCGGTGCTGCGGATCTACTCTTCACTTAAGACATC 735
Db 511 GCTTCCAGCGGCAAGTTGCTGCGCGGTGCTGCGGATCTACTCTTCACTTAAGACATC 452

Qy 736 ACGGTGGCAACGAAGCACTGGGCTGGTGGACAAAGGCGAGTACCGGATCGG 795
Db 451 ACGGTGGCAACGAAGCACTGGGCTGGTGGTGGACAAAGGCGAGTACCGGATCGG 392

Qy 796 ACCTTGTATGCCAACACCGGCAACCAAGATGAGCTTCAGGCTTCACCATCTTGCTGTC 855
Db 391 ACCTTGTATGCCAACACCGGCAACCAAGATGAGCTTCAGGCTTCACCATCTTGCTGTC 332

Qy 856 AAGGAGGTGACGAAGTTGGCTGCAAGTCTTCACTCAAGGCAAGCGGCTCTTCTAT 915
Db 331 AAGGAGGTGACGAAGTTGGCTGCAAGTCTTCACTCAAGGCAAGCGGCTCTTCTAT 272

Qy 916 GACCTTACTGAGACAGACAGCTCTTTTACGGGCTTCCATCTTATGCGAGCAGAGATGAC 975
Db 271 GACCTTACTGAGACAGACAGCTCTTTTACGGGCTTCCATCTTATGCGAGCAGAGATGAC 212

Qy 976 CCAACAGAGTATAGACATGCAACGCGGTCTTCACAGGAGGGAACAAGCTTCTGGAATT 1035
Db 211 CCAACAGAGTATAGACATGCAACGCGGTCTTCACAGGAGGGAACAAGCTTCTGGAATT 152

Qy 1036 GGGGTTTACAGAGCAAGACCCCAACTGTAGGCTGGGGGTGGGGGTGGAGTGAACGGTT 1095
Db 151 GGGGTTTACAGAGCAAGACCCCAACTGTAGGCTGGGGGTGGGGGTGGAGTGAACGGTT 92

Qy 1096 CTAGGCTCAAGGCTCAAGCTCTCGGCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1155

```

```

Db 91 CTAGGCTCAAGGCTCAAGCTCTCGGCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 32
Qy 1156 TATTCA 1161
Db 31 TATTCA 26

RESULT 10
ADP94916
ID ADP94916 standard; cDNA; 1114 BP.
XX
AC ADP94916;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human gene 11-derived ACRP30-like cDNA, SEQ ID NO:12.
XX
KW Human; ACRP30-like; adipocyte complement related protein-30-like;
KW complement factor C1q homologue; cerebellin homologue;
KW hibernation-specific protein homologue; TNF-alpha homologue;
KW tumour necrosis factor-alpha homologue; TNF-alpha homologue;
KW drug screening; diabetes mellitus; type 2 diabetes; insulin resistance;
KW obesity; metabolic disorder; endocrine disorder; antidiabetic; anorectic;
KW gene therapy; chromosome 11q11; gene; ss.
XX
OS Homo sapiens.
XX
PN M02003031586-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002MO-US032432.
XX
PR 12-OCT-2001; 2001US-0328419P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (BLON/) BLONDEL O.
PA (RUBE/) RUBEN S M.
XX
PI Blondel O, Ruben SM;
XX
DR WPI; 2003-381707/36.
XX
DR P-PSDB; ADP94857.
XX
PT New ACRP30-like nucleic acid molecules and polypeptides, useful for
PT preventing, treating or ameliorating a medical condition, such as insulin
PT resistance or diabetes mellitus.
XX
PS Claim 1, SEQ ID NO 12; 602pp; English.
XX
XX
The invention relates to novel human adipocyte complement related protein
30 (ACRP30)-like cDNAs derived from 20 genes, and their encoded proteins.
The secretion of ACRP30 is promoted by insulin, and ACRP30 acts in a
similar manner to insulin, such as reducing blood glucose levels and
reducing levels of plasma free fatty acids. ACRP30 is homologous to
complement factor C1q, cerebellin, and hibernation-specific proteins HP-
20, HP-25, and HP-27, and has a C-terminal globular domain similar in
crystal structure to C1q and tumour necrosis factor-alpha (TNF-alpha);
the latter of which has been implicated in insulin resistance in obesity
and type 2 diabetes. Additionally ACRP30 is highly expressed in
adipocytes during adipogenesis, and may also have immune-related
functions. ACRP30-like proteins may therefore be useful in the treatment
of insulin-related disorders such as diabetes and obesity. The invention
also relates to recombinant vectors and host cells comprising an ACRP30-
like nucleic acid; the recombinant production of ACRP30-like proteins;
antibodies specific for ACRP30-like proteins; use of ACRP30-like nucleic
acids or proteins in preventing, treating or ameliorating a medical
condition in an individual; methods of diagnosing a pathological
condition, or a susceptibility to a pathological condition; methods of
screening for binding partners of ACRP30-like proteins; and methods of
screening for modulators of ACRP30-like activity. The ACRP30-like nucleic
acids and proteins, and methods of the invention are useful for
preventing, treating or ameliorating a medical condition such as diabetes

```





```

QY 119 CGAGGTAACCAACCATGATCCCTGGGTGCTCTGGCTGTGACCTCCCTGCTGCTG 178
DB 1 CGAGGTAACCAACCATGATCCCTGGGTGCTCTGGCTGTGACCTCCCTGCTGCTG 60
QY 179 ACCCACTGCTTGGGCGCTTTGCTGCGAGGACTTCCGAAAAGGCTCCCTCAACTGGTCT 238
DB 61 ACCCACTGCTTGGGCGCTTTGCTGCGAGGACTTCCGAAAAGGCTCCCTCAACTGGTCT 120
QY 239 GCGAGCTGCTGGGCGCCGAGGGCCACCGGGCCCGGAGAGGCCCGAGGGCCCTCAGGAA 298
DB 121 GCGAGCTGCTGGGCGCCGAGGGCCACCGGGCCCGGAGAGGCCCGAGGGCCCTCAGGAA 180
QY 299 TGAATGGAGCAATAGGAGCTTTCTGTCGAAAGACGCGCAAGATGACACGACCGGACCGG 358
DB 181 TGAATGGAGCAATAGGAGCTTTCTGTCGAAAGACGCGCAAGATGACACGACCGGACCGG 240
QY 359 GGGACAGCGGAGAGGAAAGTTCACCTTGCCGGAACAGGTAAACCGGGGAAAGCAGAGCCAA 418
DB 241 GGGACAGCGGAGAGGAAAGTTCACCTTGCCGGAACAGGTAAACCGGGGAAAGCAGAGCCAA 300
QY 419 AGGGCAAGCGGGGCGCATTTGGGGGGCTGGCCCGCTGGCGCCGAGGGGGTCAACGGTA 478
DB 301 AGGGCAAGCGGGGCGCATTTGGGGGGCTGGCCCGCTGGCGCCGAGGGGGTCAACGGTA 360
QY 479 CCCCAGGAAAGCATGAGCAACAGGCAAGAGGGGCCAAGAGGAGGAGGAGGAGGAGGAGG 538
DB 361 CCCCAGGAAAGCATGAGCAACAGGCAAGAGGGGCCAAGAGGAGGAGGAGGAGGAGGAGG 420
QY 539 GCTTCCAGGCGCTTGGAGCTGTGAGTGGCCATACCAAGTCAAGCTTTCTCGGTGGAG 598
DB 421 GCTTCCAGGCGCTTGGAGCTGTGAGTGGCCATACCAAGTCAAGCTTTCTCGGTGGAG 480
QY 599 TGAACCAAGAGTACCAAGGAGGAGGCGTGGCCCATCAAGTTGAAGAATCTGATGAGACG 658
DB 481 TGAACCAAGAGTACCAAGGAGGAGGCGTGGCCCATCAAGTTGAAGAATCTGATGAGACG 540
QY 659 AGGGTGGCCACTCAATGCTTCCAGCGGCAAGTTCCTGCGGCGTGTGCTGGAGTCTACT 718
DB 541 AGGGTGGCCACTCAATGCTTCCAGCGGCAAGTTCCTGCGGCGTGTGCTGGAGTCTACT 600
QY 719 ACTTCACTTACGATCATCGCTGGCGCAACAGCACTTGCCCATATGGCTGTGTCACAGC 778
DB 601 ACTTCACTTACGATCATCGCTGGCGCAACAGCACTTGCCCATATGGCTGTGTCACAGC 660
QY 779 GCCAGTACCGGATCCGAGCTTTGATGTCACACCGGCAACCAAGATGTGGCTCAGGCT 838
DB 661 GCCAGTACCGGATCCGAGCTTTGATGTCACACCGGCAACCAAGATGTGGCTCAGGCT 720
QY 839 CCAACCATCTGGCTCTCAAGCAGGGGTGACGAAAGTTGGCTGACAGATCTTCTACTCAGAGC 898
DB 721 CCAACCATCTGGCTCTCAAGCAGGGGTGACGAAAGTTGGCTGACAGATCTTCTACTCAGAGC 780
QY 899 AGAAGCGGCTTTCTTATGACCTTCTACTGAGACAGACGCTTTTACGGGCTTCTCTAATCT 958
DB 781 AGAAGCGGCTTTCTTATGACCTTCTACTGAGACAGACGCTTTTACGGGCTTCTCTAATCT 840
QY 959 ATGCGGACCAAGATGACCCGACAGAGGATATGACATGCAAGGGGCTCTCAGAGGAGG 1018
DB 841 ATGCGGACCAAGATGACCCGACAGAGGATATGACATGCAAGGGGCTCTCAGAGGAGG 900
QY 1019 AACCAAGCTTCTGACTTGGGCTTACAGAGCAA 1050
DB 901 AACCAAGCTTCTGACTTGGGCTTACAGAGCAA 932

```

RESULT 13  
ID ACC42469  
AC ACC42469 standard; DNA; 885 BP.

XX ACC42469;  
XX 26-AUG-2003 (first entry)  
XX

```

DE Murine C1QSF2/5 coding sequence.
XX
XX Murine; antiinflammatory; immunosuppressive; antipeptidic; thymomimetic;
XX cytoostatic; antimicrobial; antiasthmatic; C1QSF3; C1QSF; cancer;
XX Tumour Necrosis Factor; TNF; Complement C1q; inflammatory disorder;
XX Immune disorder; autoimmune disorder; infection; cell adhesion; asthma;
XX prolatias; Hashimoto's thyroiditis; inflammatory bowel disease; gene;
XX Crohn's disease; cell activity; cell behaviour; cell development; ds.
XX
XX Mus musculus.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..885
XX FT /*tag= a
XX FT /product= "C1QSF2/5"
XX
XX W02003022992-A2.
XX
XX 20-MAR-2003.
XX
XX 04-SEP-2002; 2002MO-US027974.
XX
XX 06-SEP-2001; 2001US-0317988P.
XX
XX (SCHE ) SCHERRING CORP.
XX
XX Basham BE, Foreythe IJ, Gorman DM, Mattson JD, Moshrefi M;
XX Parham C;
XX WPI, 2003-313239/30.
XX
XX P-PSDB; ABP70818.
XX
XX Novel isolated human or murine C1QSF3 polypeptide, and antibodies
XX specific to the polypeptides, which are useful for treating prolatias,
XX metastasis, Hashimoto's thyroiditis, asthma, or inflammatory bowel
XX disease.
XX
XX Disclosure; Page 47-48; 97pp; English.
XX
XX
XX The present invention relates to human and murine C1QSF3 proteins
XX (ABP70819 and ABP70820) and their coding sequences (ACC42470 and
XX ACC42471). C1QSF proteins are members of the Tumour Necrosis Factor (TNF)
XX and Complement C1q family. The C1QSF proteins are useful for treating a
XX subject suffering from inflammatory disorder, immune disorder or
XX autoimmune disorder. The C1QSF proteins are also useful for treating
XX infection, abnormal proliferation, cancer, metastasis, or pathological
XX cell adhesion. The C1QSF proteins are also useful for treating prolatias,
XX Hashimoto's thyroiditis, asthma, or inflammatory bowel disease, e.g.,
XX Crohn's disease. The C1QSF proteins are also useful for modulating the
XX activity, behaviour, and development of cells, including monocytes, mast
XX cells, dendritic cells, macrophages, lymphocytes or neutrophils. The
XX CC protein encoded by the present sequence was used in a sequence alignment
XX with the C1QSF3 proteins
XX
XX
XX Sequence 885 BP; 208 A; 253 C; 262 G; 162 T; 0 U; 0 Other;
XX
XX
XX Query Match 60.6%; Score 703.6; DB 8; Length 885;
XX Best Local Similarity 88.0%; Pred. No. 2e-155;
XX Matches 766; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
XX
XX
XX 121 AGGGTAACCAACCATGATCCCTGGGTGCTCTGGCTGTGACCTCCCTGCTGCTGAC 180
XX 16 AAGGTGACCAACATGATCTCTGGGTACTCTTGGCTGTGCTCCCTTCCGTGTGCTGAC 75
DB
XX 181 CCACTGCTTGGCGCTTTGCTGCGAGGACTTCCGAAAAGGCTCCCTCAACTGGTTCG 240
DB 76 CCAATGCTTGGTGTCTTGTCTCCGAGGACTTCCAGAAAGGGGGTCTCAACTAGTTTC 135
QY 241 AGCTGCTGGGCGCCGAGGCGCAACCGGCGCCCGCAGAGAGCCCGAGGGCTCTCAGAGATG 300
DB 136 AGCTGCTGGTGTCTCCAGAGGCGCACTGGCGCTTCCAGAGACACAGGATCTCAGAGATG 195
QY 301 ATGGGACGAATGGGCTTCTCTGGCAAGAGCGGCCAAGATGAGACAGAGCGGACCGGGGG 360

```



```

Db      196 GTGGGAATAATGGTTCCCTCGGAAAGACCGCCAAAGTGGCAGAGACGAGACCGGGGG 255
Qy      361 GACAGCGGAGAGAAAGTCCACTGGCGGAGAGTAAACCGGGAAACCGAGCAAAAG 420
Db      256 GACAGTGAAGAAAGTCCACTGGCAGAGACCGGCAACCGTGGAAACCAAGAGCAAAAG 315
Qy      421 GCGAAAGCCGGGCGCATTTGGGCGGGCTGGCCCCCTGGCCCCCAAGGGGTCAAGCTTAC 480
Db      316 GCGAAAGCTGGGGCGATTGGGAGAGCTGGCCCTCAGAGACCCAGGGGGTCAAGTGTACC 375
Qy      481 CCCGGGAAGCATGCGACACAGGCGCAAGAGGGGCCCAAGGGGCAAGAAAGGAGCCAGGC 540
Db      376 CCGGGGAAGCATGCGACACCGGGCAAGAGGAGCCTTAAGGGCAAAAAGGGAGCTTGGC 435
Qy      541 CTCCTCAGGCCCCCTGAGCTGTGGCAGTGGCGCATACCAAGTCAAGCTTCTCGTGGCAGTG 600
Db      436 CTCCTCAGGCCCCCTGAGCTGTGGCAGTGGCGCATACCAAGTCAAGCTTCTCGTGGCAGTG 495
Qy      601 ACCAAGAGCTACCCACCGGAGCGGTGGCCCATCAAGTTTGACAAAGATTCTGATGAACGAG 660
Db      496 ACCAAGAGCTACCCACCGGAGAGACTGTGCTATCAAGTTTGACAAAGATTCTGATGAACGAG 555
Qy      661 GGTGGCCACTCAATGCTTCCAGCGGCAAGTTGCTGTGGCGGTGGCTGGATCTACTAC 720
Db      556 GGTGGCCACTCAACGCTCCAGTGGCAAGTTGCTGTGGCGGTGGCTGGATCTACTAC 615
Qy      721 TTCACTCAAGCATCAGCTGGCCCAACAGCACTGGCCATCGGCTGGTGGCAACAGGC 780
Db      616 TTTCATCATGACATTAAGCTGGCCCAACAGCACTGGCCATCGGCTGGTGGCAACATGCT 675
Qy      781 CAGTACCGCATCCGGAACCTTTGATGACCAACACCGGCAACCAAGATGTGGCTCAGGCTCC 840
Db      676 CAGTACCGCATTCGACCTTTGATGACCAACACCGGCAACCAAGATGTGGCTCAGGCTCC 735
Qy      841 ACCATCTGGCTCTCAAGCAGAGGTGACGAAGTTGGCTGCAAGATCTTCACTCAGAGCAG 900
Db      736 ACCATCTGGCTCTCAAGCAGAGGTGATGAAGTCTGGCTGCAAGATCTTCACTCAGAGCAG 795
Qy      901 AACGGGCTCTTCTATGACCTTTACTGACAGACAGCTCTTTACGGGGCTTCTTAATCTAT 960
Db      796 AATGGCTCTTCTATGACCTTTACTGACAGACAGCTCTTTACCGGGCTTCTTAATCTAT 855
Qy      961 GCCGACCAAGATGACCCCAAGAGGTATAG 990
Db      856 GCTGACCAAGAGACCCCAAGAGGTATAG 885

```

## RESULT 14

AA261744 standard; cDNA; 1107 BP.

AA261744;

27-MAR-2000 (first entry)

cDNA encoding rat skin cell secreted protein, SEQ ID NO:217.

Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;  
 embryonic skin cell; keratinocyte stem cell; transit amplifying cell;  
 secreted; transmembrane; inflammation; cancer; neurological disease;  
 angiogenesis; tumour vascularisation; growth disorder;  
 developmental disorder; skin wound; hair follicle disorder;  
 anti-inflammatory; cyostatic; neuroprotective; vulnery; ss.

Rattus sp.

MO9955865-A1.

04-NOV-1999.

29-APR-1999; 99WO-N2000051.

XX

```

PR      29-APR-1998; 98US-00069726.
PR      09-NOV-1998; 98US-00188930.
PA      (GENE-) GENESIS RES & DEV CORP LTD.
PI      Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
XX      WPI, 2000-072177/06.
DR      P-PSDB; AAY76039.
XX      Novel polynucleotides useful for the treatment of various conditions
PT      including wounds and cancer.
XX      Claim 1; Page 142-143; 235pp; English.
XX      The invention relates to novel nucleic acid sequences derived from rat
CC      dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC      and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC      cells. Polypeptides of the invention may be used to treat inflammation,
CC      cancer and neurological diseases. The proteins may be used to stimulate
CC      the growth and motility of keratinocytes, to inhibit the growth of cancer
CC      cells, to modulate angiogenesis and tumour vascularisation, to modulate
CC      skin inflammation, to modulate epithelial cell growth and to inhibit
CC      binding of HIV-1 to leukocytes. The invention may also be used to treat
CC      growth and developmental defects, skin wounds and hair follicle
CC      disorders. Sequences AA261606-261832 represent cDNA sequences derived
CC      from several mouse, rat or human skin cell types. Sequences AA261606-
CC      261649, AA261725-261765, AA261802-261811 and AA261826 encode proteins
CC      with an N-terminal signal sequence, indicating that the proteins are
CC      secreted. Sequences AA261650-261668, AA261766-261780, AA261812-261817 and
CC      AA261827-261829 encode proteins with one or more putative transmembrane
CC      domains
XX      SQ
XX      Sequence 1107 BP; 273 A; 298 C; 328 G; 208 T; 0 U; 0 Other;
SQ      Query Match      60.5%; Score 702.2; DB 3; Length 1107;
        Best Local Similarity 84.7%; Pred. No. 4.5e-155;
        Matches 800; Conservative 0; Mismatches 143; Indels 2; Gaps 1;

```

Db	642	ACCAAGATTACCAACGTAGAGCATGCGCCATCAAGTTGACAAATTCATGATGATAG	701
Qy	661	GGTGGCCACTACAAATGCTTTCCAGCGGCAAGTTGCTCGCGGTGCTGGAGATCTATAC	720
Db	702	GGAGGCGCACTCAATATGCATCCAGTGGCAAGTTGCTGCTCAGCGTGCACGGAGATCTATATAC	761
Qy	721	TTTACCTATCGACATGACGGCTGGCCAAACAAGCACCTGGCCATCCGGCTGCTGATGCAACCGGC	780
Db	762	TTTACCTATGACATTTACGCTGGCCAAACAACCTGGCCATCCGGCTGCTGATGCAACATGGC	821
Qy	781	CAGTACCGCATCCGACCTTTGATGTCCAACCCGGCAACCAAGATGAGGCTCAGAGTCC	840
Db	822	CAGTACCGCATTCGGACTTTTGACGCCAACCCGCAACCAAGAGTGGGCTCGGCTCC	881
Qy	841	ACCATCTCTGGCTCTTAACGAGGGGTACGAAAGTTTGGCTGCAGATTTCTATCTCAAGACGAG	900
Db	882	ACCATCTCTAGGCTCTTAACGAGGGGTATTAAGTCTGGTTACAGATTTTCTATCTCGAAGCAG	941
Qy	901	AACGGGCTCTCTATGACGCTTACTACGACAGACTTTTACGGGGCTTCTCTAATCTAT	960
Db	942	AATGGACTCTTCTATGACACCTTTATTGGACCGACAGCTGTTCACCGGCTTCTCATCTAC	1001
Qy	961	GCCGACCAAGATGACCCCAACGAGGTATAGACATGCCACGG--CGGTCTTCAGGACGAGG	1018
Db	1002	GCTGATCAAGAGAGACCCCAATGAGGTTATGACAACTGGGGGTGAGCGCTGTCCAGCAGGG	1061
Qy	1019	AACCAAGCTTCTGGACTTGGGCTTACAGACCAAGACCCCAACATCG	1063
Db	1062	ACTAAGATTCGCAAGGGGTGCTGATGAGAGAGGATCTCTGAATCTG	1106

RESULT 15  
AAC99677  
ID AAC99677 standard; cDNA; 1107 BP.

AAC99677;

DT	08-MAR-2001	(first entry)
----	-------------	---------------

DE Skin cell cDNA, SEQ ID NO: 217.

KW Rat; skin cell; cytostatic; anti-inflammatory; anti-HIV; nootropic;

KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;

XX

XX

XX

XX

[illegible]

XX

XX

PI Murison JG;

DR WPI; 2001-007495/01.

XX

PT disorders and encoding polypeptides used for treating inflammatory

XXXXXX

XX

CC mammalian skin cells. The polypeptide is useful for stimulating

modulating angiogenesis, inhibiting angiogenesis and vascularisation of tumours, modulating skin inflammation, stimulating the growth of epithelial cells, inhibiting the binding of human immunodeficiency virus (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and neurological diseases. The polynucleotide can be used as a marker, in the identification of genetic disorders, and for the design of oligonucleotides for examining expression patterns

Query Match	60.5%	Score 702.2;	DB 4;	Length 1107;
Best Local Similarity	84.7%;	Pred. No. 4.5e-155;		
Matches 800; Conservative	0;	Mismatches 143;	Indels 2;	Gaps 1;

QY	12	AGGGTAACCAACAGATATCCCTGGGATGCTCTGGACTATGCTCCCTGATGCTGTAC	180
Db	162	AAGGTGACCAACATGATCTCTGGATGCTCTTGGGCTGTGCTCCCTTGCTGTAC	221
QY	181	CCACTGCTTGGGCGCTTTTGTCTGTGCAAGGACTTTCGGAAAGGCTCCCTCAACTGTGCTGC	240
Db	222	CCAAATGCTTGGTGCTTTGTGCTGTGAGGGGACTTTCAGAAAGGGGTGGTCCCACTGGTGTGTC	281
QY	241	AGCTGCTGTGGCCCCCGAGGGCCACCCGGCCCCCGAGAGGCCCGAGGGCCCTCAGGATG	300
Db	282	AGTCTGCTGTATCCCAAGGCCACCTGGCCCTCCAGAGCAACAGATTCCTCAGGAATG	341
QY	301	ATGGGACGAAATGGGCTTTCCTTGGCAAAAGCCGACAGATGACACGACGGCGACCGGGG	360
Db	342	GTTGGGAAGAAATGGGTTTTCCTGTGTAAGGATGGCCAAAGCCGCCAGGACCGAGACCGAGG	401
QY	361	GACACGCGAGAGGAAGTTCACCTTGGCCCGGACAGGTAAACCGGGGAAGCCAGACCAAG	420
Db	402	GACAGTGAAGAAAGGTGTCACTTGGCAGGACAGGCAACGAGGAAACAAAGGACCAAG	461
QY	421	GGCAAAAGCCGGGGCCATTGGGCGGGGCTTGCCCCGTGGCCCCCAAGGGGGTCAACGTTAC	480
Db	462	GGCAAAAGCTGGGGCCATTGGGAGAGCCGGTCTCTGAGGACCCCAAGGGGGTCAAGTGTAC	521
QY	481	CCCGGAAAGCATGSCACACAGGCAAGAGGGGGCCAAAGGCAAGAAAGGGAGCCAGGC	540
Db	522	CCCGGAAACATGTGTTATCCCGGCAAGAGGGGACCTTAAGGGCAAGAAAGGGGAACTGGG	581
QY	541	CTCCAGGGCCCTGACAGCTGTGGGAGTGGGCAATACAAAGTACAGCTTCTCGGTGGCAATG	600
Db	582	CTCCAGGGCCCTGTAGTGTGGGCAATGAGCCAGAGCAAGTGGGCTTTTGGGTGGCGTTA	641
QY	601	ACCAAGAGCTACCAAGCGGAGCGGCTGGCCATCAAGTTTGAACAAGTTCTGATGAACGAG	660
Db	642	ACCAAGAGTTATCCACGTGAGCGACTGCCCCATCAAGTTTGAACAAGTTCTGATGAATAG	701
QY	661	GGTGGCCACTACATGCTTTCAGAGGGCAAGTTTGTCTGCGGCGTGGCTGGGATCTACATAC	720
Db	702	GGAGGCCACTACATGCAATCAAGTGGCAAGTTTGTCTGACGCTGGCCAGGGAATCTAATTAC	761
QY	721	TTCAACTACGACATACGCTGGSCCAACAAGACCTGGSCATTCGSCCTGTGTGCAACAAGC	780
Db	762	TTTACCTATGACATTTACGCTGGCCAAACAACACTGGGCAATGGGCTATGTGACATATGGC	821
QY	781	CAGTACCGGATCCGGACTTTGATGTCCACAACCCGGCAACAACGATGTGGCTCAGGCTTC	840
Db	822	CAGTACCGCAATTCGGACTTTTGAAGCCCAACCCGGCAACACGACTGTGGCTTCGGGCTTC	881
QY	841	ACCATTCCTGGGCTCCAGCAGGAGTGAAGAGTTTGGCTGCGAGATCTTGTACTCAGAGAG	900
Db	882	ACCATTCCTAGCTCTCAGAGAGGAGTGAATGATCTGTGTTACAGATTTTCTACTCGGAGAG	941
QY	901	AACGGGCTCTTCTATGACCTTTACTGTGACAGACAGGCTTTTACGGGCTTCTAATCTAT	960
Db	942	AATGGAGCTCTTCTAGACCTTTATTTGGACGAGCAGGCTGTTACAGGGCTTCTCATCTAC	1001
QY	961	GCCGACCAAGATGACCCCAAGAGGTATAGCAATGCAACGG--GGGTCTCTCAGGCAAGG	1011
Db	1002	GCTATCAAGAGAGCCCAATGAGGTTATAGCAAGCTGGGGTGTGAGCGGTCCAGGCAAGG	1061

Qy 1019 AACAGCTTCTGCACTTGGGCTTACAGACAGACCCCACTG 1063  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1062 ACTAAGATTCCGCAAGGGTGGCTGATAGAGAGATCTCTGAAC TG 1106

Search completed: December 25, 2004, 04:01:21  
Job time : 663 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2004, 02:34:02 ; Search time 4236 Seconds

(without alignments)  
9987.377 Million cell updates/sec

Title: US-10-621-787-1

1161

Perfect score: 1 ggaactactgctgctgggccc.....atcccaacttttattca 1161

Sequence:

IDENTITY\_NUC

Scoring table: Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0\*

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_est4:\*

5: gb\_est5:\*

6: gb\_est6:\*

7: gb\_est7:\*

8: gb\_est8:\*

9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	779.4	67.1	863	5	B0881020 AGENCOURT
2	716.2	61.7	1196	3	AK007683 Mus muscu
3	645.2	55.6	671	5	BQ006349 UI-H-E11-
4	630.4	54.3	654	4	BM675363 UI-E-E10-
5	607.8	52.4	806	4	B1109819 602901018
6	597	51.4	634	4	BM714600 UI-E-E10-
7	588.2	50.7	708	4	BI489428 603020916
8	547	47.1	714	6	CF108213 Shultzomi
9	535.4	46.1	544	1	AI916779 wd26e01.x
10	523.8	45.1	535	2	AM304481 xv22h08.x
11	500.6	43.1	659	4	BI822472 603038061
12	493	42.5	557	1	AA732948 2978d10.8
13	466	40.1	904	4	BI488418 603020916
14	458.8	39.5	526	4	BG381862 296753 MA
15	452.4	39.0	455	1	AI288625 q188b05.x
16	438.6	37.8	496	2	BE232317 137306 MA
17	432	37.2	559	1	AV616838 AV616838
18	413.8	35.6	539	5	BQ207443 UI-R-DY1-
19	413.4	35.6	556	2	AM965223 ESTJ77296
20	398.6	34.3	444	1	AA443244 aa14c05.x
21	397.6	34.2	696	5	BQ210972 UI-R-DY1-
22	392.8	33.8	715	7	CK76809 968175 MA
23	388.6	33.5	670	6	CB324225 UI-R-DY0-
24	387.4	33.4	826	7	CK794745 AGENCOURT

25	383.2	33.0	663	6	CB323171	CB323171 UI-R-DY0-
26	373.2	32.1	426	1	AA194106	AA194106 ZR18903.8
27	369.6	31.8	528	4	BG670534	BG670534 DRNBCC10
28	356	30.7	449	8	A2805724	A2805724 ZM0067H16
29	338.4	29.1	409	7	W79527	W79527 zd81e04.81
30	333	28.7	351	1	A1038644	A1038644 ox39d11.81
31	326.8	28.1	544	1	AA443206	AA443206 aa14c06.8
32	326.8	28.1	564	2	BE851196	BE851196 ua93d01.Y
33	311.8	26.9	374	7	W41045	W41045 mc35b09.r1
34	307.4	26.5	870	9	AY420023	AY420023 Mus muscu
35	305.8	26.3	1436	9	AK078818	AK078818 Mus muscu
36	304.4	26.2	463	6	CB735431	CB735431 AMGNNUC:N
37	295.4	25.4	870	9	AY420021	AY420021 Homo sapi
38	286.8	24.7	502	4	BI303579	BI303579 UI-R-DY0-
39	284	24.5	451	6	CB787179	CB787179 AMGNNUC:N
40	274.6	23.7	583	4	BI696475	BI696475 603347416
41	272.4	23.5	483	5	BX527335	BX527335 BX527335
42	269.6	23.2	870	9	AY420022	AY420022 Pan trogl
43	264.8	22.8	293	2	BF511014	BF511014 UI-H-B14-
44	264.8	22.7	714	5	BU127923	BU127923 603114908
45	263.2	22.7	734	5	BU323601	BU323601 603492824

#### ALIGNMENTS

RESULT 1  
B0881020  
LOCUS  
DEFINITION  
AGENCOURT 8124216 Lupaki dorsal root ganglion Homo sapiens cDNA  
clone IMAGE:6178402 5', mRNA sequence.  
B0881020  
VERSION  
B0881020.1 GI:22273028  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Nih-MGC http://mgc.nci.nih.gov/  
1 (bases 1 to 863)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strauberg, Ph.D.  
Email: cgabs-remail.nih.gov

Tissue Procurement: Dr. James R. Lupski  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1AM13557 row: 0 column: 11  
High quality sequence stop: 722.  
Location/Qualifiers

#### FEATURES

##### SOURCE

1. 863  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6178402"  
/sex="male"  
/tissue\_type="dorsal root ganglia"  
/dev\_stage="adult", 36 yr"  
/lab\_host="DH10B"  
/clone\_lib="Lupaki dorsal root ganglion"  
/note="Vector: PCMV-SPORE (Life Technologies); Site 1:  
NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCACCAAGCGCTCCG-3' and  
5'-GACTAGTTTGAATCGGAGCGGCCCTT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life

ORIGIN Technologies."

Query Match 67.1%; Score 779.4; DB 5; Length 863;  
Best Local Similarity 97.3%; Pred. No. 2.4e-177;  
Matches 825; Conservative 0; Mismatches 17; Indels 6; Gaps 3;

24 GCTGTGCGCGGCTGTGCGCGGTGAGAAAGCGGAGCGCGCGCGAGAGCTTCT 83  
Db 1 GCTGTGCGCGGCTGTGCGCGGTGAGAAAGCGGAGCGCGCGCGAGAGCTTCT 60  
QY 84 TTGCTCGGAGCGCGCTGTGAGAGTGGCGGCGAGCGGAGTAAACAATGATCCCTG 143  
Db 61 TTGCTCGGAGCGCGCTGTGAGAGTGGCGGCGAGCGGAGTAAACAATGATCCCTG 120  
QY 144 GTGCTCTGTGCGCGCTGTGCGCGGTGAGAGTGGCGGAGTAAACAATGATCCCTG 203  
Db 121 GTGCTCTGTGCGCGCTGTGCGCGGTGAGAGTGGCGGAGTAAACAATGATCCCTG 180  
QY 204 CAGGAGACTTCGCGAAAGGCTCCCTCAACTGTGTGAGCTGTGAGCTGTGAGCGCC 263  
Db 181 CAGGAGACTTCGCGAAAGGCTCCCTCAACTGTGTGAGCTGTGAGCGCC 240  
QY 264 ACCCGGCGCGCGCGAGCGCGCGCGCTCAAGATGATGAGAGTGGCTTTCTG 323  
Db 241 ACCCGGCGCGCGCGAGCGCGCGCGCTCAAGATGATGAGAGTGGCTTTCTG 300  
QY 324 CAAAGAGCGCGCAATGAGAGCGAGCGCGCGCGCGCGCGAGAGAGTCCACC 383  
Db 301 CAAAGAGCGCGCAATGAGAGCGAGCGCGCGCGCGCGAGAGAGTCCACC 360  
QY 384 TGCGCGGAGAGTAAACGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 443  
Db 361 TGCGCGGAGAGTAAACGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
QY 444 GGTGCG 503  
Db 421 GGTGCG 480  
QY 504 CAAAGAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 563  
Db 481 CAAAGAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
QY 564 CAGTGGCGCATACCAAGTCAAGTCTGTGCGTGGAGTGAACCAAGAGTACCAAGAG 623  
Db 541 CAGTGGCGCATACCAAGTCAAGTCTGTGCGTGGAGTGAACCAAGAGTACCAAGAG 600  
QY 624 GGTGCGCGCATACCAAGTCTGTGCGTGGAGTGAACCAAGAGTACCAAGTCTG 683  
Db 601 GGTGCGCGCATACCAAGTCTGTGCGTGGAGTGAACCAAGAGTACCAAGTCTG 660  
QY 684 CCGGAAATGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 743  
Db 661 CCGGAAATGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720  
QY 744 CAAAGAGCGCGCATACCAAGTCTGTGCGTGGAGTGAACCAAGAGTACCAAGAG 802  
Db 721 CAAAGAGCGCGCATACCAAGTCTGTGCGTGGAGTGAACCAAGAGTACCAAGAG 780  
QY 803 ATGCGCAACACCGG---CAACCAAGATGTGCGCTTCAGGCTTCACAT-CTGCTCTCAA 857  
Db 781 ATGCGCAACACCGGCAACCAAGATGTGCGCTTCAGGCTTCACATCTCTCTCAA 840  
QY 858 GCAGGGTG 865  
Db 841 GAAAGGG 848

RESULT 2  
AK007683 1196 bp mRNA linear HTC 03-APR-2004  
LOCUS AK007683  
DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length  
enriched library, clone:1810033K05 product:COMPLEMENT-C10 TUMOR

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE

AUTHORS

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

NECROSIS FACTOR-RELATED PROTEIN (COMPLEMENT-C10 TUMOR NECROSIS  
FACTOR-RELATED PROTEIN 2) homolog [Homo sapiens], full insert  
sequence.  
AK007683 GI:12841383  
AK007683.1 GI:12841383  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Komano, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,  
Komano, H., Akiyama, J., Nishi, K., Kikunai, T., Taahiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Harada, M., Nishino, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
PANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5  
The PANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1196)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,  
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,  
Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,  
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komano, H., Kouda, M.,  
Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,  
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,  
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,  
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,  
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,  
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp,  
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
Please visit our web site (http://genome.gsc.riken.jp/) for further  
details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken



## FEATURES

### source

CDS

polyA\_signal

## ORIGIN

## Matche

Qy	121	AGGGTAAACCAACCAAGATCCCTGGGTGTCTCTGGACCTGTGGCTCCCTCTGTGCTGTGAC	180
Db	155	AAGGTGACAAACATGATCTCCTGGGTACTCTTGGCTGTGTGCTTCCTCGGTGTGCTGTAC	214
Qy	181	CCAATGCTTGGACGGCTTTTGTCTGGCAGAGGACTTCCGGAAAGGCTCCCTCTCACTGTGTCTGC	240
Db	215	CCAAATGCTTGGTGTCTTGTCTTGTGAGAGGACTTCAAGAAAGGGAGGTCTCTCAACTGTATTTGC	274
Qy	241	AGCTGTGCTGGCCCCCAGAGGCCCAACCCGGCCCCCAGAGGCCCAAGAGGCCCTCAGAGATG	300
Db	275	AGCTGTGCTGGTCTCCCAAGGCCCAACTGTGGCTCTCAAGAGGACCAAGGATCTCTCAGAGATG	334
Qy	301	ATGGGAGGAATGGGCTTCTCTGGAGAAAGCGGCCAATATGCAACAGACGGGACCGGGGG	360
Db	335	GTGGGAGGAATGGGCTTCTCTGGAGAAAGCGGCCAATATGCAAGATGGCCAGAGCGGAGACCGGGGG	394
Qy	361	GACAGCGAGAGGAAGTCCACTTGGCCGAGCAGGTAACTCGGGGAAACCAAGACCAAG	420
Db	395	GACAGTGAAGAAAGAGTCCACTTGGCAGAGCAAGCAACCTGTGAAAACAAGACCAAG	454
Qy	421	GGCAAACCGGGGGCAATTGGGGGGGCTGGGCCCTCGTGGCCCCAAGGGGGTCAACGTATTC	480
Db	455	GGCAAAGCTGGGGCCATTGGCGAGCTGGGCCCTCGAGGACCCAAAGGGGGTCAAGTGTATTC	514

QY	481	CCCGGAAAGCATGCGACACACAGGCAAGAGGGGCCCAAGGCAAGAAAGGGAGCCAGGC	540
Db	515	CCCGGAAAGCATGCGACACCGGGGAGAGAGGAGACTTAAAGGCAAAAAAGGGAGCGCTGGC	574
QY	541	CTCCAGAGCCCTGACGCTGTGGCATGGGCCATACCAAGTCAGCTTCTCCGTGGCAGTG	600
Db	575	CTCCAGAGCCCTGACGCTGTGGCAGTAGAGCCAGACCAAGTCGCGCTTCTCGTGGCGGTGA	634
QY	601	ACCAAGAGCTAACCCAGGGAGCGGCGTCGCCCATCAAGTTTGAAGAATTCGATGAAAGAG	660
Db	635	ACCAAGAGCTAACCCAGGGAGCGGCGTCGCTTCAAGTTTGAAGAATTCGATGAAAGAG	694
QY	661	GGTGGCCACTACAAATGCTTCCAGCGGAGCAAGTTGCTGTGCGGCGTGCCTGGATCTACTAC	720
Db	695	GGTGGCCACTACAAAGCGTCCAGTGGCAAGTTGCTGTGACGCGTGCCTGGATCTACTAC	754
QY	721	TTTGAACCTACGACATCAAGCTGGGCGCAACAAGACACTGGCCATCGGGCTGTGGTCAAAAGGC	780
Db	755	TTTGAACCTACGACATTAAGCTGGGCGCAACAACTGGCCATCGGCTGTGGTCAAAAGGT	814
QY	781	CAGTACCGCATCCCGGACCTTTGATGGCCACACCGGCAACACGATGTGGCTCAAGCTCC	840
Db	815	CAGTACCGCATCCCGGACTTTTGAATGCCAACACCGGCAACACGATGTGGCTCCGCGCTCC	874
QY	841	ACCATCTCGGCTTCAAGCAGAGGTGACGAAAGTTGGCTGCAGATCTTCTACTCAGAAGAG	900
Db	875	ACCATCTCGGCTTCAAGAGAGGGTGAAGAGTCTGGCTGCAGATCTTCTACTCAGAAGAG	934
QY	901	AAACGGGCTTTCTATGACCCCTTACTTGGACAGAGCCTCTTATAGGGGCTTCTATCTAT	960
Db	935	AATGGCCTCTTCTATGACCCCTTACTTGGACAGAGCCTCTTACTCAGGCTTCTATCTAC	994
QY	961	GCCGACCAAGATGACCCCAACGAGGTATGACATGCACGG-CCGTCCTCAGGCAAGGGA	1019
Db	995	GCTACCAAGAGAGCCCAACGAGGTATGACCAAGCCGGGGTATGAGCTTGAAGTAGGGA	1054
QY	1020	ACAAGCTTCTGGAATTGAGGCGTTACAGAGCAAGACCCCAACTGATAGGCTGGGGGTGGGG	1079
Db	1055	CTAAGAGCTG--GTTGGGTGCTGGAGAGAAAGTCCCTCACTGGGGCTGTGGAAGTGA	1112
QY	1080	GGTGAAGTGAAGCGGTTCTAGCCTCAGGCTCACTCTCCGCTCTTTTTCCTCCCTCAT	1139
Db	1113	ATCTTGGGATCTTTATTTCCAGGCAAGGCGTCTCTATTTGCTGTAAAAAAGAAATCAT	1172
QY	1140	TAAATCCAAACCTTT	1154
Db	1173	TAAATCCAAAGTATT	1187

[illegible]

through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
 Seq primer: M13 FORWARD  
 POLYA=yes

## FEATURES

Source

Location/Qualifiers

1..671  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5845954"  
 /tissue\_type="Chondrosarcoma"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_id="NCI CGAP B11"  
 /note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia) with a modified polylinker. Site 1: EcoR I; Site 2: Not I; NCI CGAP B11 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is  
 ACGTGGCAC.  
 TAG\_TISSUE=chondrosarcoma  
 TAG\_LIB=UI-H-B11  
 TAG\_SEQ=ACACTTGCAC"

## ORIGIN

Query Match 55.6%; Score 645.2; DB 5; Length 671;  
 Best Local Similarity 99.5%; Pred. No. 5.9e-145;

Matches 647; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 512 GGGCCAAAGGCAAGAAAGGAGGAGCCCTCCAGAGCCCTGAGCTGTG3CAGT9GCC 571  
 Db 671 GGGCCAAAGGCAAGAAAGGAGGAGCCCTCCAGAGCCCTGAGCTGTG3CAGT9GCC 612  
 QY 572 ATACCAAGTCAGCTTTCTCGGTGAGTACCAAGTACCCACGGAGGCGCTGCCCA 631  
 Db 611 ATACCAAGTCAGCTTTCTCGGTGAGTACCAAGTACCCACGGAGGCGCTGCCCA 552  
 QY 632 TCAAGTTGACAAATCTGATGAACGAGGCTGACCTAACAAGCTTCCAGCGGCAAGT 691  
 Db 551 TCAAGTTGACAAATCTGATGAACGAGGCTGACCTAACAAGCTTCCAGCGGCAAGT 492  
 QY 692 TCGTCTGCGGCGTCTGAGATCTACTACTTCACTAGACATCAGCGTGGCAACAAGC 751  
 Db 491 TCGTCTGCGGCGTCTGAGATCTACTACTTCACTAGACATCAGCGTGGCAACAAGC 432  
 QY 752 ACCTGGCATGGGCTGCTGTCACAAAGGCGATCCGATCCGACCTTTGATGCCAACA 811  
 Db 431 ACCTGGCATGGGCTGCTGTCACAAAGGCGATCCGATCCGACCTTTGATGCCAACA 372  
 QY 812 CCGGCAACCAAGATGGGCTCAGAGCTCCACATCTGCTCTCAAGAGGGGTGACGAG 871  
 Db 371 CCGGCAACCAAGATGGGCTCAGAGCTCCACATCTGCTCTCAAGAGGGGTGACGAG 312  
 QY 872 TTTGGCTGAGATCTTCTACTCAGAGCAAGCGGCTTTCTATGACCTTTACTGACAG 931  
 Db 311 TTTGGCTGAGATCTTCTACTCAGAGCAAGCGGCTTTCTATGACCTTTACTGACAG 252  
 QY 932 ACAGCTCTTTACGGGCTTCTAATCTATGCGGACCAAGATGACCCCAAGAGTATAGA 991  
 Db 251 ACAGCTCTTTACGGGCTTCTAATCTATGCGGACCAAGATGACCCCAAGAGTATAGA 192  
 QY 992 CATGCCACGGGCTCTTCCAGGAGGAAACAAGCTTCTGAGCTTGGGCTTACAGAGCAAG 1051  
 Db 191 CATGCCACGGGCTCTTCCAGGAGGAAACAAGCTTCTGAGCTTGGGCTTACAGAGCAAG 132  
 QY 1052 ACCCCACAACGTATGAGGCTGGGGGCTGAGTGAACGGTTCTAGCCTCAAGCTCAC 1111

Db 131 ACCCCACAACGTATGAGGCTGGGGGCTGAGTGAACGGTTCTAGCCTCAAGGCTCAC 72  
 QY 1112 CTCTCGGCTCTTTTTCCTTCCCTTCACTTAATCAACCTTTTATTTCA 1161  
 Db 71 CTCTCGGCTCTTTTTCCTTCCCTTCACTTAATCAACCTTTTATTTCA 22

## RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITL

JOURNAL

MEDLINE

PUBMED

COMMENT

|||||  
 654 bp mRNA 1linear EST 27-FEB-2002  
 UI-E-EJ0-ahr-p-10-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone  
 UI-E-EJ0-ahr-p-10-0-UI 3', mRNA sequence.  
 BM675363  
 BM675363.1 GI:18985261  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 654)  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 8889548  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.regen.com).  
 Seq primer: M13 Forward  
 POLYA=yes.

## FEATURES

Source

Location/Qualifiers

1..654  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="fetal eyes, lens, eye anterior segment,  
 optic nerve, retina, Retina Foveal and Macular, RPE and  
 Choroid"  
 /dev\_stage="fetal and adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_id="UI-E-EJ0"  
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker. Site 1: EcoR I; Site 2: Not I;  
 UI-E-EJ0 is a subtracted cDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tags for this library are: fetal eyes,  
 AARGCCGAT; optic nerve, CGATTAGCA; eye anterior segment,  
 AGATGACAGA; lens, CGATTAGCA; retina, CCGCG;  
 Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This  
 library was created for the program, Gene Discovery in the  
 Visual System, supported by National Eye Institute (NEI).  
 TAG\_TISSUE=RPE and Choroid  
 TAG\_LIB=UI-E-EJ0

## ORIGIN

TAG\_SEQ=ACCTA"

Query Match 54.3%; Score 630.4; DB 4; Length 654;  
 Best Local Similarity 99.8%; Pred. No. 2.2e-141;  
 Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

530 GGGAGCCAGGCTCCGAGCCCTGAGCTGTGGCAGGTCATACCAAGTCACTTTCT 589  
 |||  
 653 GGGAGCCAGGCTCCGAGCCCTGAGCTGTGGCAGGTCATACCAAGTCACTTTCT 594  
 |||  
 590 CGGTGAGTGAACCAAGAGTACCAAGGAGCGGCTGCCATCAAGTTTGACAAGATTC 649  
 |||  
 593 CGGTGAGTGAACCAAGAGTACCAAGGAGCGGCTGCCATCAAGTTTGACAAGATTC 534  
 |||  
 650 TGATGAACGAGGCTGCGCACTACATGCTTCAGCGGCAAGTTGCTGCGGCTGCTG 709  
 |||  
 533 TGATGAACGAGGCTGCGCACTACATGCTTCAGCGGCAAGTTGCTGCGGCTGCTG 474  
 |||  
 710 GGAATCTACTTCACTAGACATCAAGCTGGGCAACAGCACTGGCCATCGGCTG 769  
 |||  
 473 GGAATCTACTTCACTAGACATCAAGCTGGGCAACAGCACTGGCCATCGGCTG 414  
 |||  
 770 TGCAACAAGCCAGTACCGGATCCGATCCGATCTTGTGATGCAACCGGCAACGATG 829  
 |||  
 413 TGCAACAAGCCAGTACCGGATCCGATCCGATCTTGTGATGCAACCGGCAACGATG 354  
 |||  
 830 CCTGAGCTCCACCATCTGCTCTCAAGCAGGCTGACGAAGTTGGCTGCAATCTTCT 889  
 |||  
 353 CCTGAGCTCCACCATCTGCTCTCAAGCAGGCTGACGAAGTTGGCTGCAATCTTCT 294  
 |||  
 890 ACTCAGAGCAAGAGGCTCTTCTTANGACCTTACCTGAGCAGAGCTCTTTAGGGCT 949  
 |||  
 293 ACTCAGAGCAAGAGGCTCTTCTTANGACCTTACCTGAGCAGAGCTCTTTAGGGCT 234  
 |||  
 950 TCCTAATCTATGCGACCAAGATGACCCCAAGAGTATGACATGCGACGCGGCTCTC 1009  
 |||  
 233 TCCTAATCTATGCGACCAAGATGACCCCAAGAGTATGACATGCGACGCGGCTCTC 174  
 |||  
 1010 CAGGCAAGGAAACAGCTTCTGAGCTTGGGCTTACAGAGCAAGAGCCCAACCTGAGGCT 1069  
 |||  
 173 CAGGCAAGGAAACAGCTTCTGAGCTTGGGCTTACAGAGCAAGAGCCCAACCTGAGGCT 114  
 |||  
 1070 GGGGGTGGGGGCTGAGTGAAGCGGCTTACGCTCAGGCTCACCTCCGCTCTTTT 1129  
 |||  
 113 GGGGGTGGGGGCTGAGTGAAGCGGCTTACGCTCAGGCTCACCTCCGCTCTTTT 54  
 |||  
 1130 TCCCTTCATTAATCCAAACCTTTTATTC 1161  
 |||  
 53 TCCCTTCATTAATCCAAACCTTTTATTC 22  
 |||

RESULT 5  
 B1109819 806 bp mRNA linear EST 26-JUN-2001  
 LOCUS 602901018P1 NCI\_GCAP\_Mams Mus musculus cDNA clone IMAGE:5030717 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1109819  
 VERSION B1109819.1 GI:14560720  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 806)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strauberg, Ph.D.  
 Email: gspds-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
 http://image.lnl.gov  
 Plate: LMNL1086 row: C column: 06  
 High quality sequence stop: 806.  
 Location/Qualifiers

## FEATURES

source

1. 806  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="mix FVB/N, C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5030717"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="7 months"  
 /lab\_host="DH10B"  
 /clone\_id="NCI\_GCAP\_Mams5"  
 /note="Organ: mammary; Vector: pCMV-Sport6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Robin Humphreys,  
 NIH"

## ORIGIN

Query Match 52.4%; Score 607.8; DB 4; Length 806;  
 Best Local Similarity 88.0%; Pred. No. 6.7e-136;  
 Matches 696; Conservative 0; Mismatches 92; Indels 3; Gaps 3;

174 TGCTGACCACTGCTGGCGCTTGTGCTGAGGGAATTCGGAAGGCTCCCTCACT 233  
 |||  
 1 TGCTGACCACTGCTGGCGCTTGTGCTGAGGGAATTCGGAAGGCGGCTCTCACT 60  
 |||  
 234 GGTGTGAGCTGTGCTGGCCCTCCAGGGCCCAAGCCGCTCCAGAGACCCCAAGGCTC 293  
 |||  
 61 AGTTGAGGCTGTGCTGTGCTCCCAAGGCCCACTGGCTCCAGAGCAACGAGATCTC 120  
 |||  
 294 AGGAATGATGGAAGCAATGGGCTTCTCTGGCAAGAGCGCCAAAGATGACAGAGCGCA 353  
 |||  
 121 AGAGTGTGTGGAAGATGGGTTTCCCTGGAAAGAGCGCCAAAGATGGCCAGAGCGAGA 180  
 |||  
 354 CCGGGGGGACAGCGGAGAGGAAGGTCCACTGGCCGAGCAGATTAACCGGGGAAACCAAG 413  
 |||  
 181 CCGGGGGGACAGTGAAGAGAGGTCCACTGGCCGAGCAGAGCAACCGTGAAGAAACAGG 240  
 |||  
 414 ACCAAAGGCAAGACCGGGCCATTGGGCGGGCTGGCCCTCCAGAGGGGTCAA 473  
 |||  
 241 ACCAAAGGCAAGACCTGGGCTATTGGCAGAGCTGGCCCTCCAGAGCCCAAGGGGCTCAG 300  
 |||  
 474 CGGTACCCCGGGAAGATGAGCACACAGGCAAGAGGGGCCAAAGGGAAGGGA 533  
 |||  
 301 TGGTACCCCGGGAAGATGAGCACACCGGCAAGAGGACTTAAGGGCAAAAAGGGA 360  
 |||  
 534 GCCAGGCTCCAGAGCCCTGAGCTGTGCGATGAGGCATACCAAGTCAAGTTCTCGGT 593  
 |||  
 361 GCTGAGCTCCAGAGCCCTGAGCTGTGCGATGAGGCATACCAAGTCAAGTTCTCGGT 420  
 |||  
 594 GGCAGTGAACCAAGCTTACCAAGGAGCGGCTGCCATCAAGTTTGACAAGATTCGAT 653  
 |||  
 421 GGCAGTGAACCAAGCTTACCAAGGAGCGGCTGCCATCAAGTTTGACAAGATTCGAT 480  
 |||  
 654 GAAAGAGGTGGCCACTACATGCTTCCAGCGGCAAGTTGCTGCGGCGTGGGAT 713  
 |||  
 481 GAAAGAGGTGGCCACTACATGCTTCCAGCGGCAAGTTGCTGCGGCGTGGGAT 540  
 |||  
 714 CTACTACTTCACTACAGACATCAAGCTGGCCCAACAGCACTGGCCATGGCTGGTGA 773  
 |||  
 541 CTACTACTTCACTACAGACATCAAGCTGGCCCAACAGCACTGGCCATGGCTGGTGA 600  
 |||  
 774 CAAGGCAAGTACCGATCCGAGCTTGTGATGCCAACAAGGCAACCAAGATGAGGCTC 833  
 |||  
 601 CAATGTGCAATACCGATCCGAGCTTGTGATGCCAACAAGGCAACCAAGATGAGGCTC 660  
 |||  
 834 AGGCTCACCATCTGCTCTCAA-GCAGGAGTGAAGAGTTGGCTGCAAGTCTTCTACT 892  
 |||

```

|||||
Db 661 GGCTCCACCATCTAGCTCTCAAGGCGAGGTGTAAGTCTGGCTGCAGATCTTACT 720
QY 893 CAGAGGAGACGGGCTTTCTATGACCCCTTACGAGACAGACGCTTT-TAGGGGTTG 951
Db 721 CAGAGCAGAAATGGCTCTTCTTACGA-CCTTACTGACCGACGCTGTCTAACCGGCTTC 779
QY 952 CTATCTATGC 962
Db 780 CTATCTATGC 790

RESULT 6
BM714600 634 bp mRNA linear EST 28-FEB-2002
LOCUS UI-E-EJ0-ahr-p-10-0-UI-r1 UI-E-EJ0 Homo sapiens cDNA clone
DEFINITION UI-E-EJ0-ahr-p-10-0-UI-r1 UI-E-EJ0 Homo sapiens cDNA clone
ACCESSION BM714600
VERSION BM714600.1 GI:19027858
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 634)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
JOURNAL MEDLINE
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..634
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahr-p-10-0-UI-r1"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPS and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1lb="UI-E-EJ0"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-EJ0 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (drr)18 tail. The
sequence tags for this library are: fetal eyes,
AGATCAAGA; lens, CGATTAGCA; eye anterior segment,
AATGCCGAT; optic nerve, CCATTAGTG; retina, CCGCC; Retina
Foveal and Macular, GTCC; RPS and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the

```

```

Visual System, supported by National Eye Institute (NEI)."
ORIGIN
Query Match 51.4%; Score 597, DB 4; Length 634;
Best Local Similarity 99.0%; Pred. No. 2.6e-133;
Matches 611; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 2 GAAACTATGCTCGGGGCGGACGCTCTGCGCGCTGCGCTGAGGAAAGCGGAGC 61
Db 1 GAAGACTATGCTCGGGGCGGACGCTCTGCGCGCTGCGCTGAGGAAAGCGGAGC 60
QY 62 CGAGGCCCGCGGAGGCTTTCTTGTCTCGGAGCGCCCTTGAGAGTGCGGGCGAGCGG 121
Db 61 CGAGGCCCGCGGAGGCTTTCTTGTCTCGGAGCGCCCTTGAGAGTGCGGGCGAGCGG 120
QY 122 GGGTAAACCAAGATATCCCTGGGATGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTG 181
Db 121 GGGTAAACCAAGATATCCCTGGGATGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTG 180
QY 182 CACTGTGTGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAACTGTGTGCA 241
Db 181 CACTGTGTGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAACTGTGTGCA 240
QY 242 GCTGCTGCGCTTGTCTGCGAGGAGCTTCCGAGGAGCTTCCGAGGAGCTTCCGAGG 301
Db 241 GCTGCTGCGCTTGTCTGCGAGGAGCTTCCGAGGAGCTTCCGAGGAGCTTCCGAGG 300
QY 302 TGGAGCAATGGGCTTCTTCTGCGAAAGCGGCAAGTGAACAGCGGAGCGGAGG 361
Db 301 TGGAGCAATGGGCTTCTTCTGCGAAAGCGGCAAGTGAACAGCGGAGCGGAGG 360
QY 362 ACAAGCGAGAGGAGGTCACCTTGTGCGAGCAAGTAAACCGGAGAAAGCCAGCAAG 421
Db 361 ACAAGCGAGAGGAGGTCACCTTGTGCGAGCAAGTAAACCGGAGAAAGCCAGCAAG 420
QY 422 GCAAGCGGAGGAGGTCACCTTGTGCGAGGAGGTCACCTTGTGCGAGGAGGAGG 481
Db 421 GCAAGCGGAGGAGGTCACCTTGTGCGAGGAGGTCACCTTGTGCGAGGAGGAGG 480
QY 482 CGGAGAGGAGGTCACCTTGTGCGAGGAGGTCACCTTGTGCGAGGAGGAGGAGG 541
Db 481 CGGAGAGGAGGTCACCTTGTGCGAGGAGGTCACCTTGTGCGAGGAGGAGGAGG 540
QY 542 TCCAGAGGAGGTCACCTTGTGCGAGGAGGTCACCTTGTGCGAGGAGGAGGAGG 601
Db 541 TCCAGAGGAGGTCACCTTGTGCGAGGAGGTCACCTTGTGCGAGGAGGAGGAGG 599
QY 602 CCAAGAGCTACCGACG 618
Db 600 CCAAGAGCTACCGACG 616

RESULT 7
BI489428/c 708 bp mRNA linear EST 28-AUG-2001
LOCUS BI489428
DEFINITION 603020916T1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:519189 3',
mRNA sequence.
ACCESSION BI489428
VERSION BI489428.1 GI:15328656
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 708)
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
REFERENCE Unpublished (1999)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Contact: Robert Strausberg, Ph.D.
JOURNAL Email: cgabs-remail.nih.gov
COMMENT Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

```



Db 303 TATATCCGGGCAAGAGGACCTAAGGGCAAGAAAGGGAACTGGGCTCCCAAGGCCCTG 362  
 QY 555 CAGCTGTGGAGTGGCCATACCAAGTCACTTCTCGGTGGCACTGACCAAGACTACCC 614  
 Db 363 TAGCTGGGAGTAGGCGAGCGAAGCTGGCTTTTCGTGGCGGTAAACAAAGATTACCC 422  
 QY 615 ACCGAGACGGCTGCGCCATCAAGTTTGAACAGATTCTATGAAACAGAGGTGGCCACTCAA 674  
 Db 423 ACCGAGACGACTGCGCCATCAAGTTTGAACAGATTCTATGAAAGAGGAGCCACTCAA 482  
 QY 675 TGCCTTCAGCGGCAAGTTTGGTCTGCGCGGTGGCTGGAGATTCTACTTCACTTACGACAT 734  
 Db 483 TGCATCAGTGGCAAGTTTGGTCTGCGCGGTGGCTGGAGATTCTACTTCACTTACGACAT 542  
 QY 735 CAGCGTGGCCAAACAAGACCTGGCCATCGCGCTGGTGCACAAGCGGCAATGCGCATCG 794  
 Db 543 TAGCGTGGCCAAACAAGACCTGGCCATCGCGCTGGTGCACAAGCGGCAATGCGCATCG 602  
 QY 795 GACCTTGTATGACCAACACCGGCAACACAGATGGCTTCCAGCTTCTGCTCT 854  
 Db 603 GACTTTGAGCGCCAAACACCGGCAACACAGATGGCTTCCAGCTTCTGCTCT 662  
 QY 855 CAAGCAGGTGACGAGATTGGCTGCAAGATCTTCTACTCAGAGCAAGACG 905  
 Db 663 CAAGCAGGTGACGAGATTGGCTGCAAGATCTTCTACTCAGAGCAAGATG 713

RESULT 9 544 bp mRNA linear EST 17-DEC-1999  
 AI916779/c w26e01.x1 NCI CGAP G66 Homo sapiens cDNA clone IMAGE:2106808 3' similar to SW:ACR3 HUMAN Q15848 30 KD ADIPOCYTE COMPLEMENT-RELATED PROTEIN PRECURSOR 7, mRNA sequence.

ACCESSION AI916779 GI:5636634  
 VERSION AI916779.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 544)  
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
 www.bio.lnl.gov/bbrp/image/image.html  
 Insert length: 1061 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 454.  
 Location/Qualifiers

## FEATURES

source 1..544  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2106808"  
 /tissue\_type="pooled germ cell tumors"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP G66"  
 /notes="Vector: p7773D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI CGAP G64 was prepared, and as circles were made in vitro. Following HAP purification,

## ORIGIN

Query Match 46.1%; Score 535.4; DB 1; Length 544;  
 Best Local Similarity 99.8%; Pred. No. 1.9e-118;  
 Matches 536; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1257096-1258631, 1469064-1470983, and 1475592-1476743).  
 Subtraction by Bento Soares and M. Fatima Bonaldo.

QY 625 CTGCCATCAAGTTTGAACAGATTCTGATGAAGAGGGTGGCCACTAAGATGCTTCCAGC 684  
 Db 544 CTGCCATCAAGTTTGAACAGATTCTGATGAAGAGGGTGGCCACTAAGATGCTTCCAGC 485  
 QY 685 GCGAAGTTGCTGCGCGCGGCTGGGATCTACTACTTCACTACCAAGCAATCAGCTGGCC 744  
 Db 484 GCGAAGTTGCTGCGCGCGGCTGGGATCTACTACTTCACTACCAAGCAATCAGCTGGCC 425  
 QY 745 AACAGACCTGGCATCGGCTGTGACAAAGCCAGTACCGCATCCGGAACCTTGTAT 804  
 Db 424 AACAGACCTGGCATCGGCTGTGACAAAGCCAGTACCGCATCCGGAACCTTGTAT 365  
 QY 805 GCCAACAACCGGCAACCAAGTGGCTTCAAGCTTCCACATCTGCTTCAAGCAGGT 864  
 Db 364 GCCAACAACCGGCAACCAAGTGGCTTCAAGCTTCCACATCTGCTTCAAGCAGGT 305  
 QY 865 GACGAAGTTGGCTGACATCTTCTACTGACAGACAGAGGGCTTCTATGACCTTAC 924  
 Db 304 GACGAAGTTGGCTGACATCTTCTACTGACAGACAGAGGGCTTCTATGACCTTAC 245  
 QY 925 TGACAGACAGCTCTTTAAGGGCTTCTATGATCCAGCAGATGAGTGAAGCCCAAGAG 984  
 Db 244 TGACAGACAGCTCTTTAAGGGCTTCTATGATCCAGCAGATGAGTGAAGCCCAAGAG 185  
 QY 985 GTATGACATGCAACCGGCTCTTCCAGGCAAGGAAACAGCTTGGACTTGGGCTTACA 1044  
 Db 184 GTATGACATGCAACCGGCTCTTCCAGGCAAGGAAACAGCTTGGACTTGGGCTTACA 125  
 QY 1045 GACCAAGACCCCAACACTGTAAGGCTGGGGGTGGAGTGAAGGGCTTACAGCTCA 1104  
 Db 124 GACCAAGACCCCAACACTGTAAGGCTGGGGGTGGAGTGAAGGGCTTACAGCTCA 65  
 QY 1105 GGCTCACTCTCTCGGCTCTTTTTCCTGATTAATCAACCTTTTATTC 1161  
 Db 64 GGCTCACTCTCTCGGCTCTTTTTCCTGATTAATCAACCTTTTATTC 8

RESULT 10 535 bp mRNA linear EST 18-JAN-2000  
 AM304481/c xv22h08.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone IMAGE:2813919 3' similar to SW:ACR3 HUMAN Q15848 30 KD ADIPOCYTE COMPLEMENT-RELATED PROTEIN PRECURSOR 7, mRNA sequence.

ACCESSION AM304481 GI:6714170  
 VERSION AM304481  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 535)  
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 This clone is available royalty-free through LINL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 410.  
 Location/Qualifiers

FEATURES



source

1. .535  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2813919"  
/lab\_host="DH10B"  
/clone\_1pb="Soares NFL T GBC S1"  
/notes="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBH1.9W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The diver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bernaldo."

ORIGIN

Query Match 45.1%; Score 523.8; DB 2; Length 535;  
Best Local Similarity 98.7%; Pred. No. 1.2e-115;  
Matches 528; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 625 CTGCCCCATAGTTGACAAATCTGATGACGAGGCTGCGCACTAATGCTTCCAGC 684  
DB 535 CTGCCCCATAGTTGACAAATCTGATGACGAGGCTGCGCACTAATGCTTCCATC 476  
QY 685 GGCAGTTCCTGCGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 744  
DB 475 GCCAAGTTCCTGCGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 416  
QY 745 AACAGACCTGCGCTGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 804  
DB 415 AACAGACCTGCGCTGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 356  
QY 805 GCCAAGCCTGCGCTGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 864  
DB 355 GCCAAGCCTGCGCTGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 296  
QY 865 GACGAATTTGGCTGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 924  
DB 295 GACGAATTTGGCTGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 236  
QY 925 TGGACAGACGCTCTTTACGCGCTTCTAATCTATGCGGACAGATGACCCCAAGAG 984  
DB 235 TGGACAGACGCTCTTTACGCGCTTCTAATCTATGCGGACAGATGACCCCAAGAG 176  
QY 985 GTATAGCATGCGCGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 1044  
DB 175 GTATAGCATGCGCGCGCTGCGCTGAGATCTACTTCACTGACATCAGCTGCGC 116  
QY 1045 GAGCAAGACCCCAACATGATGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1104  
DB 115 GAGCAAGACCCCAACATGATGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 56  
QY 1105 GGGTCACCTCTCTGCGCTCTTTTCCCTTCAATTAATCAAACTTTTATT 1159  
DB 55 GGGTCACCTCTCTGCGCTCTTTTCCCTTCAATTAATCAAACTTTTATT 1

RESULT 11  
BI822472 659 bp mRNA linear EST 04-OCT-2001  
LOCUS 603038061.F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5179095 5',  
DEFINITION mRNA sequence.  
ACCESSION BI822472.1 GI:15934022  
VERSION BI822472.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS NIH-MGC  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strauberg, Ph.D.  
Email: rgs@nihs.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
plate: LNL1146 row: 1 column: 16  
High quality sequence stop: 659.  
Location/Qualifiers

FEATURES  
source

1. .659  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5179095"  
/lab\_host="DH10B"  
/clone\_1pb="NIH MGC 115"  
/notes="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site 1: Not 1; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (score site is destroyed upon cloning). Average insert size 1.8 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 43.1%; Score 500.6; DB 4; Length 659;  
Best Local Similarity 95.0%; Pred. No. 5.1e-110;  
Matches 626; Conservative 0; Mismatches 19; Indels 14; Gaps 10;

QY 20 CGAGCTCTGCGCGCTGCTGCGCTGAGGAAAGCGGAGCCCGCGG-AGAG 78  
DB 1 CGAGCTCTGCGCGCTGCTGCGCTGAGGAAAGCGGAGCCCGCGGAGAG 60  
QY 79 CTTCTTTGCTCCGAGCGCCCTGAGCGTGGCGGCGAGCGGAGGTAACCATATATC 138  
DB 61 CTTCTTTGCTCCGAGCGCCCTGAGCGTGGCGGCGAGGTAACCATATATC 120  
QY 139 CCGTGGGTCCTGCGCTGCGCTGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 197  
DB 121 CCGTGGGTCCTGCGCTGCGCTGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
QY 198 TGCTCGAGGAGCTTCCGAAAGGCTTCCC--TCACTGCTGCGAGCTGCGGCGCC 255  
DB 181 TGCTCGAGGAGCTTCCGAAAGGCTTCCCCTGTAAGCTGCTGCGAGCTGCGGCGCC 240  
QY 256 CAGGCGCCACCC-GGCCCCCAGAGGCGGAGGCGCTCAGAAATG--TGGCAAAATG 312  
DB 241 CAGGCGCCACCCCGGCGCCCGCAGAGGCGGAGGCGCTCAGAAATGAGTGGAGAAATG 300  
QY 313 GGGCTTCCTGGAAGAGCGGCAAGATGACAGAGGCGGAGGAGGAGGAGGAGGAGGAG 372  
DB 301 GGGCTTCCTGGAAGAGCGGCAAGATGACAGAGGCGGAGGAGGAGGAGGAGGAGGAG 360  
QY 373 GAAAGTCACCTGCGCGGAGGTAACCGGAGGAAAGCAGAGCCAA--GGGCAAAAGCGG 431  
DB 361 GAAAGTCACCTGCGCGGAGGTAACCGGAGGAAAGCAGAGCCAAATGGGCAAAAGCGG 420  
QY 432 GGCATTGGCGGCGGCTGGCCCGGTGGCCCA--GGGGGTCAACGGTAACCCCGGAAAGC 490  
DB 421 GGCATTGGCGGCGGCTGGCCCGGTGGCCCAAGGGGGGTCAACGGTAACCCCGGAAAGC 480  
QY 491 ATGGCAACACGAGGCAAGAGGCGGCAAGGAGGCAAGAAAGGAGGAGGAGGAGGAGGAG 550



```

/clon="IMAGE:5191898"
/lab host="DH10B"
/clone.lib="NIH_MGC_114"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

```

```

ORIGIN
Query Match      40.1%; Score 466; DB 4; Length 904;
Best Local Similarity 77.3%; Pred. No. 1.2e-101;
Matches 782; Conservative 0; Mismatches 10; Indels 220; Gaps 9;

QY 1 GGAAGACTATGCTGGGGGCGAGCGCTGCGCGGCTGCGCGTGAAGGCGGAGC 60
DB 84 GGAAGACTATGCTGGGGGCGAGCGCTGCGCGGCTGCGCGTGAAGGCGGAGC 143
QY 61 GCGAGGCGCGCGAGAGCTTCTTGTCTCCGAGCGCCCTGAGAGTGGCGGAGCGCG 120
DB 144 GCGAGGCGCGCGAGAGCTTCTTGTCTCCGAGCGCCCTGAGAGTGGCGGAGCGCG 203
QY 121 AGGGTAACACATGATCCCTGGGAGTCTCTGCGCTGTGCTCCCTGTGCTGAC 180
DB 204 AGGGTAACACATGATCCCTGGGAGTCTCTGCGCTGTGCTCCCTGTGCTGAC 263
QY 181 CCACTGCTGGGCGCTTGTCTGCGAGGAGACTCCGGAAGGCTCCCTCACTGTCTGC 240
DB 264 CCACTGCTGGGCGCTTGTCTGCGAGGAGACTCCGGAAGGCTCCCTCACTGTCTGC 286
QY 241 AGCTGCTGGGCGCTTGTCTGCGAGGAGACTCCGGAAGGCTCCCTCACTGTCTGC 300
DB 287 AGCTGCTGGGCGCTTGTCTGCGAGGAGACTCCGGAAGGCTCCCTCACTGTCTGC 303
QY 301 ATGGGAGCAATGGGCTTCTGTGCAAGACGCGCAAGTGAACAGCGCGAGCGGAGG 360
DB 304 ATGGGAGCAATGGGCTTCTGTGCAAGACGCGCAAGTGAACAGCGCGAGCGGAGG 352
QY 361 GACAGCGGAGAGGAGGTCCACTGTGGCGGAGCAGGTAAACGGGGAAGCAAG 420
DB 353 GACAGCGGAGAGGAGGTCCACTGTGGCGGAGCAGGTAAACGGGGAAGCAAG 352
QY 421 GGCAGAGCCGGGCGCATTTGGGCGGCTGGCCCCGCTGGCCCCAAGGGGTCAAGTACC 480
DB 353 GGCAGAGCCGGGCGCATTTGGGCGGCTGGCCCCGCTGGCCCCAAGGGGTCAAGTACC 352
QY 481 CCCGGAAGATGGCAGACACAGGAGAGGGGCCAAGGGCAAGAA-AGGGAGCCAGG 539
DB 353 -ACGGGAAGATGGCAGACACAGGAGAGGGGCCAAGGGCAAGAAAGGGGAGCCAGG 411
QY 540 CTTTCCCGAGGCGCTGTGCAAGTGTGGCAGTACCAAGTCACTTTCTGTGGTCAAGT 599
DB 412 CTTTCCCGAGGCGCTGTGCAAGTGTGGCAGTACCAAGTCACTTTCTGTGGTCAAGT 471
QY 600 GACCAAGAGCTAACCCAGGAGGAGCGGCTGCCCATCAAGTTTGAAGAATTCATGAAGCA 659
DB 472 GACCAAGAGCTAACCCAGGAGGAGCGGCTGCCCATCAAGTTTGAAGAATTCATGAAGCA 531
QY 660 GGGTGGCACTACATGCTTTCAGAGCGCAAGTTGCTGTGCGGCGTGGGATCTACTA 719
DB 532 GGGTGGCACTACATGCTTTCAGAGCGCAAGTTGCTGTGCGGCGTGGGATCTACTA 591
QY 720 CTTTCACTAAGACATGAGCTGTGGCAAGCAAGCACTTGGGCTGTGGTCAAG-AAAG 778
DB 592 CTTTCACTAAGACATGAGCTGTGGCAAGCAAGCACTTGGGCTGTGGTCAAG-AAAG 651
QY 779 GCGAGTACCGGATCCGAGCTTTGA-TGCCAAGACCGGCAACAGATGTGGCTCAGAGC 837
DB 652 GCGAGTACCGGATCCGAGCTTTGA-TGCCAAGACCGGCAACAGATGTGGCTCAGAGC 711

```

```

QY 838 TCCACATCTCTGCTCTCAAGCAGAGGTGACAGAGTTTGGC-TGCAGATCTTCTACTGAG- 895
DB 712 TCCACATCTCTGCTCTCAAGCAGAGGTGACAGAGTTTGGC-TGCAGATCTTCTACTGAG 771
QY 896 -AGCAGAACGGGCTCTTCTAAGACC-TTACTGACAGACAGCTCTTTACGGGCTTCT 953
DB 772 AGCAGAACGGGCTCTTCTAAGACC-TTACTGACAGACAGCTCTTTACGGGCTTCT 831
QY 954 AATTTATGCGACCAAGAG-TGACCCCAACGAGGTATGACATGCGCAGCGCG 1004
DB 832 AATTTATGCGACCAAGAGTTGACATGACCCCAACGAGGTATGACATGCGCAGCGCG 883

```

```

RESULT 14
BG381862      526 bp  mRNA  linear  EST 12-MAR-2001
LOCUS        296753 MARC 1Pig Sus scrofa cDNA 5', mRNA sequence.
DEFINITION   BG381862
ACCESSION    BG381862
VERSION      BG381862.1 GI:13306334
KEYWORDS     EST.
SOURCE       Sus scrofa (pig)
ORGANISM     Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE    1 (bases 1 to 526)
AUTHORS      Fahrentz,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Petrea,G., Sultana,R.,
Quackenbush,J. and Keele,J.W.
Pig gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
JOURNAL      MEDLINE
PUBMED       12213789
COMMENT      Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4356
Fax: 402 762 4390
Email: smitht@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGAAGACGCTATGACCAT
BACKWARD: GTTTCAGTCACGACG
Plate: 1 row: 1 column: 12
Seq primer: ATTAGGAGACACTATAG.
Location/Qualifiers
1..526

```

```

FEATURES
SOURCE
1..526
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_id="MARC 1Pig"
/notes="Vector: pCMV SPORT6; site 1: NotI; site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

```

```

ORIGIN
Query Match      39.5%; Score 458.8; DB 4; Length 526;
Best Local Similarity 92.0%; Pred. No. 6.2e-100;
Matches 484; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 216 GAAAGGCTCCCTCACTGCTGACGCTGCGCGGCGCCGAGGCGCCAGCCGCGCC 275
DB 1 GAAAGGATCCCTCACTGCTGACGCTGCGCGGCGCCGAGGCGCCAGCCGCGCC 60
QY 276 AGAGCCCGAGGCGCTCAGAGATGAGAGAGATGGGCTTTCTGCAAGAGCGGCA 335
DB 61 AGAGCCCGAGGCGCTCAGAGATGAGAGAGATGGGCTTTCTGCAAGAGATGGGCA 120

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2004, 03:17:57 ; Search time 123 Seconds

(without alignments)  
6709.154 Million cell updates/sec

Title: US-10-621-787-1

Perfect score: 1161  
Sequence: 1 ggaactactgctgctgagcc.....atccaactttttatcca 1161

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1161	100.0	1161	4	US-09-552-204A-1
2	702.2	60.5	1107	3	US-09-188-930-217
3	702.2	60.5	1107	4	US-09-312-283C-217
4	701.6	60.4	1052	4	US-09-312-283C-358
5	613.6	52.9	855	4	US-09-552-204A-10
6	453	39.0	750	3	US-09-188-930-19
7	289.6	24.9	536	4	US-09-312-283C-19
8	253.2	21.8	1333	3	US-09-552-204A-11
9	101.2	8.7	729	3	US-09-227-357-51
10	101.2	8.7	729	3	US-09-140-804-10
11	101.2	8.7	729	3	US-09-686-838B-10
12	96.8	8.3	4517	3	US-09-140-804-9
13	96.8	8.3	4517	3	US-09-686-838B-9
14	96.8	8.3	4517	4	US-09-776-976-5
15	96.8	8.3	4517	4	US-09-909-547-5
16	96.8	8.3	4545	4	US-09-569-852B-5
17	95.8	8.3	1313	3	US-08-463-911-6
18	91.6	7.9	1347	3	US-09-140-804-1
19	91.6	7.9	1347	4	US-09-686-838B-1
20	91.6	7.9	1377	4	US-09-866-028-41
21	91.6	7.9	1377	4	US-10-140-002-361
22	91.6	7.9	1377	4	US-09-944-457-41
23	89.2	7.7	20966	4	US-09-776-976-7
24	89.2	7.7	20966	4	US-09-909-547-7
25	89.2	7.7	20966	4	US-09-569-852B-1
26	87.2	7.5	1276	2	US-08-463-911-1
27	87.2	7.5	1276	4	US-09-776-976-3

28	87.2	7.5	1276	4	US-09-909-547-3	Sequence 3, Appl1
29	84.6	7.3	728	4	US-09-336-536-2	Sequence 2, Appl1
30	84.6	7.3	1338	4	US-09-336-536-1	Sequence 1, Appl1
31	83.4	7.2	2543	1	US-08-555-669-11	Sequence 11, Appl1
32	83.4	7.2	2543	3	US-09-073-663-11	Sequence 11, Appl1
33	81.4	7.0	1152	4	US-09-776-976-1	Sequence 1, Appl1
34	81.4	7.0	1152	4	US-09-909-547-1	Sequence 1, Appl1
35	80.6	6.9	4167	4	US-09-169-768-12	Sequence 12, Appl1
36	79	6.8	660	4	US-09-169-768-43	Sequence 43, Appl1
37	79	6.8	3170	4	US-09-169-768-1	Sequence 1, Appl1
38	79	6.8	3171	4	US-09-169-768-15	Sequence 15, Appl1
39	79	6.8	3171	4	US-09-169-768-19	Sequence 19, Appl1
40	79	6.8	3181	1	US-08-655-086-1	Sequence 1, Appl1
41	79	6.8	3349	4	US-09-169-768-13	Sequence 13, Appl1
42	79	6.8	3531	4	US-09-169-768-7	Sequence 7, Appl1
43	79	6.8	3541	4	US-09-169-768-9	Sequence 9, Appl1
44	79	6.8	4409	4	US-09-331-347C-22	Sequence 22, Appl1
45	78.4	6.8	1839	1	US-08-383-744-1	Sequence 1, Appl1

#### ALIGNMENTS

```
RESULT 1
US-09-552-204A-1
Sequence 1, Application US/09552204A
Patent No. 6620309
GENERAL INFORMATION:
APPLICANT: Piddington, Christopher S.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGY ZACRNP2
FILE REFERENCE: 99-08
CURRENT APPLICATION NUMBER: US/09/552,204A
CURRENT FILING DATE: 2000-04-19
PRIOR APPLICATION NUMBER: 60/130,207
PRIOR FILING DATE: 1999-04-20
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1161
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (133)...(987)
US-09-552-204A-1

Query Match 100.0%; Score 1161; DB 4; Length 1161;
Best Local Similarity 100.0%; Pred. No. 8.4e-289;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATACTATGCTGGGCGCGACGCTGTGCGGCTGCTGCGGAGGAAAGCCGGGAC 60
DB 1 GGAATACTATGCTGGGCGCGACGCTGTGCGGCTGCTGCGGAGGAAAGCCGGGAC 60
QY 61 GCGAGGCCCGCGGAGGCTTCTTGTCTCGGAGAGCCCTGGAAGTGGCGGCGCG 120
DB 61 GCGAGGCCCGCGGAGGCTTCTTGTCTCGGAGAGCCCTGGAAGTGGCGGCGCG 120
QY 121 AGGTAACACCATGATCCCTGAGGTGCTCTGTGCTGTGCGCTGCTGCTGCTGAC 180
DB 121 AGGTAACACCATGATCCCTGAGGTGCTCTGTGCTGTGCGCTGCTGCTGCTGAC 180
QY 181 CCACTGCTTGCGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAACTGCTGC 240
DB 181 CCACTGCTTGCGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAACTGCTGC 240
QY 241 AGCTGCTTGCGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAAGATG 300
DB 241 AGCTGCTTGCGCGCTTGTCTGCGAGGAGCTTCCGAAAGGCTCCCTCAAGATG 300
QY 301 ATGGAGCAATGGGCTTCTGCAAGAGCGCCAAATGAGACGAGCGGCGG 360
DB 301 ATGGAGCAATGGGCTTCTGCAAGAGCGCCAAATGAGACGAGCGGCGG 360
```

```

Db 301 ATGGAGCAATGAGGCTTTCTCTGCAAAAGACGCGCAAGATGACACAGACGGACCGGGGG 360
Qy 361 GACAGCGGAGAGGAGAGGCTCCACTGCGCGGACAGGTAAACGGGGGAAAGCCGAGCCAAAG 420
Db 361 GACAGCGGAGAGGAGAGGCTCCACTGCGCGGACAGGTAAACGGGGGAAAGCCGAGCCAAAG 420
Qy 421 GGCAGAACCGGGGGGCTTTGGGGCGGGCTGGCCCGGTGGCCCAAGGGGGGTCAACGGTACC 480
Db 421 GGCAGAACCGGGGGGCTTTGGGGCGGGCTGGCCCGGTGGCCCAAGGGGGGTCAACGGTACC 480
Qy 481 CCCGGAGAGCATGACACACAGGCAAGAGGGGCCCAAGGGGCAAGAAAGGGAGCCAGGC 540
Db 481 CCCGGAGAGCATGACACACAGGCAAGAGGGGCCCAAGGGGCAAGAAAGGGAGCCAGGC 540
Qy 541 CTCCAGAGGCGCTGACGTGTGGAGTGGCCATCAAGTCAAGCTTTCTCGGTGGCAAGT 600
Db 541 CTCCAGAGGCGCTGACGTGTGGAGTGGCCATCAAGTCAAGCTTTCTCGGTGGCAAGT 600
Qy 601 ACCAGAGCTACCCACGCGGAGCGGCTGCCATCAAGTTTGACAAGATTTGATGAACGAG 660
Db 601 ACCAGAGCTACCCACGCGGAGCGGCTGCCATCAAGTTTGACAAGATTTGATGAACGAG 660
Qy 661 GGTGGCCATCAATGCTTCCAGCGGCAAGTTGCTTGGGGCGTGGCTGGAGTTACTAC 720
Db 661 GGTGGCCATCAATGCTTCCAGCGGCAAGTTGCTTGGGGCGTGGCTGGAGTTACTAC 720
Qy 721 TTCACTACGACATCAAGCTGGCCAAAGACCTGGCCATCGGGCTGGTGAACAAGGC 780
Db 721 TTCACTACGACATCAAGCTGGCCAAAGACCTGGCCATCGGGCTGGTGAACAAGGC 780
Qy 781 CAGTACCGCATCCGAGCTTTGATGACCAACCGGCAACACAGATGGCTCAGGCTCC 840
Db 781 CAGTACCGCATCCGAGCTTTGATGACCAACCGGCAACACAGATGGCTCAGGCTCC 840
Qy 841 ACCATCTGCTCTCAAGACAGGCTGACGAGTGTGGCTGACAGATCTTCACTCAGACAG 900
Db 841 ACCATCTGCTCTCAAGACAGGCTGACGAGTGTGGCTGACAGATCTTCACTCAGACAG 900
Qy 901 AAGCGGCTCTTCAATGACCTTCACTGACAGACGCTCTTCAAGGGCTCTCAATGAT 960
Db 901 AAGCGGCTCTTCAATGACCTTCACTGACAGACGCTCTTCAAGGGCTCTCAATGAT 960
Qy 961 GCCGACAGATGACCCCAACGAGATATGACATGACGCGGCTCTCAGACGAGGAA 1020
Db 961 GCCGACAGATGACCCCAACGAGATATGACATGACGCGGCTCTCAGACGAGGAA 1020
Qy 1021 CAACTTCTGACCTTGGGCTTTACAGACGAGACCCCAACTGTAGGCTGGGGGTGGGG 1080
Db 1021 CAACTTCTGACCTTGGGCTTTACAGACGAGACCCCAACTGTAGGCTGGGGGTGGGG 1080
Qy 1081 GTGAGTGAAGGGTCTTAGGCTCAGGCTCAGCTCGGCTCTTTTTCCTTCAAT 1140
Db 1081 GTGAGTGAAGGGTCTTAGGCTCAGGCTCAGCTCGGCTCTTTTTCCTTCAAT 1140
Qy 1141 AATCAAACTTTTATTC 1161
Db 1141 AATCAAACTTTTATTC 1161

```

```

RESULT 2
US-09-188-930-217
; Sequence 217, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A

```

```

; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 217
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Rat
US-09-188-930-217

Query Match      60.5%; Score 702.2; DB 3; Length 1107;
Best Local Similarity 84.7%; Pred. No. 6.1e-171;
Matches 800; Conservative 0; Mismatches 143; Indels 2; Gaps 1;

Qy 121 AGGGTAAACACATGATCCCTGAGTGTCTCTGAGCTGCTCCCTGCTGCTGAC 180
Db 162 AAGGTAAACACATGATCTCTGATGCTCTTGGCTGTGCTCTTCCGTGTGCTGAC 221
Qy 181 CAATGCTTGGGCGCTTGTGCTGCAAGGAACTTCGGAAGGCTCCCTCAACTGTGTGC 240
Db 222 CAATGCTTGGTCTTGTGCTGCAAGGAACTTCGGAAGGCTGCTCAACTGTGTGC 281
Qy 241 AGCTGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 282 AGCTGCTGGTCTTCCCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 341
Qy 301 ATGGAGCAATGAGGCTTTCTTGGCAAGACGCGCAAGATGACACGAGCGGCGG 360
Db 342 GTGGAGAGATGAGGTTTTCTGTTAGATGCGCAAGACGCGCAAGACGAGACGAGG 401
Qy 361 GACAGCGAGAGAGAGGCTCACTGCGGCAAGTAAACGGGAAAGCAAGCAAG 420
Db 402 GACAGTGAAGAGAGGCTCACTGCGGCAAGGCAAGGCAAGGAAAGCAAGCAAG 461
Qy 421 GGCAGAGCGGGGCGATTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Db 462 GGCAGAGCTGGGCGATTTGGGAGAGCGGCTTCAGAGACCCAGGCGGCTCAGTACC 521
Qy 481 CCCGGAAGATGAGCACACGAGCAAGAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGC 540
Db 522 CCCGGAAGATGATGTATACCGGCGAGAGAGGACCTTAGGGCAAGAGAGGAACTGGG 581
Qy 541 CTCCAGAGGCGCTGACGCTGTGCAAGTGGCCATACCAAGTCACTTCTCGGTGGCAGTG 600
Db 582 CTCCAGAGGCGCTGTAGCTGCGGCAAGTGGCCAGGCAAGTGGCTTCTCGGTGGCGTGA 641
Qy 601 ACCAGAGCTACCCACGAGGAGCGGCTGCCATCAAGTTTGAAGAAATTCAGTGAAGAG 660
Db 642 ACCAGAGTTACCCACGAGGAGCGGCTGCATCAAGTTTGAAGAAATTCAGTGAAGAG 701
Qy 661 GGTGGCCATCAATGCTTCAAGCGGCAAGTTGCTGCGGCGGCTGCGGATCTACTAC 720
Db 702 GAGAGCCATCAATGATGATCAAGGCAAGTTGCTGAGCGGCGGCGGAGGATTAATAC 761
Qy 721 TTCACTACGACATCAAGCTTGGCCCAACAGACCTGGCCATTCGGCTGTGGCACAGGC 780
Db 762 TTTAATATGACATTAAGCTGGCCCAACAAACACTGGCCATTCGGCTGTGGCACAGGC 821
Qy 781 CAGTACCGCATTCGGAAGCTTTGATGCGCAACGCGGCAACAGATGGGCTCAGGCTCC 840
Db 822 CAGTACCGCATTCGGAAGCTTTGATGCGCAACGCGGCAACAGATGGGCTCAGGCTCC 881
Qy 841 ACCATCTGCTCTCAAGACAGGATGACGAGTGTGGCTGACAGATCTTCACTACAGACAG 900
Db 882 ACCATCTGCTCTCAAGAGAGGATGAGTGTGGTTAAGATTTCTACTGGAGACAG 941
Qy 901 AAGCGGCTCTTCAATGACCTTCACTGACAGACGCTCTTTACGCGGCTTCTAATATAT 960
Db 942 AATGACTTCTTCAAGACCTTATTTAGACGAGCCTGTTCACCGGCTTCTCATATAC 1001
Qy 961 GCCGACAGATGACCCCAACGAGATATGACATGCCACG--CGGTCTCCAGGCGAGG 1018
Db 1002 GCTGATCAAGAGACCCCAATGAGATATGACAGAGCTGGGGTGGAGCGGTCCAGGAGG 1061

```





```
QY 421 GCGAAGCCGCGGCAATTGGCGGCTGCGCCCGCTGGCCCAAGGGGCTCAACGGTACC 480
Db 333 GCGAAGCTGGGCGCAATTGGGAGAGGGGGTCTCGAGGACCCAAAGGGGGTCACTGGTACC 392
QY 481 CCGGGAGCATGGCACACCAAGGCGAAGAGGGGCCCAAGGGCAAGAAAGGGGACCGAGC 540
Db 393 CCGGGAGCATGGTATACCGGGCAAGAGGAGCTTAAGGGCAAGAAAGGGGAACTGGG 452
QY 541 CTCGAGGCGCCGCGAGCTGTGGCACTGAGGCGATACCAAGTCACTTCTCGTGGGAGTG 600
Db 453 CTCGAGGCGCCCTGTAGCTGTGGCGAGTACCGAGCCAAAGTGGCTTTTCGGTGGGTA 512
QY 601 ACCAAGAGCTACCCACGAGGCGGCTGCCCATCAAGTTTGAACAAGATTCTGATGAAGAG 660
Db 513 ACCAAGAGTTTACCAAGTGAAGAGTGGCCATCAAGTTTGAACAAGATTCTGATGAAGAG 572
QY 661 GGTGGCACTACAAATGCTTCCAGCGGCAAGTTGCTGCGCGGCTGCTGGGAATCTATAC 720
Db 573 GGAAGGCACTACAAATGCTTCCAGTGGCAAGTTGCTGCGCGGCTGCGAGGATCTATAC 632
QY 721 TTCACTACGACATCAAGCTGAGCAACAGCACTGGGCAATGGGCTGTGGCAACAAGGCG 780
Db 633 TTCACTATGACATTAAGCTGAGCAACAGCACTGGGCAATGGGCTGTGGCAACAATGGC 652
QY 781 CAGTACCGGATCCGGAACCTTTGATGCGCAACACCGGCAACGATGTGGCTGAGGCTCC 840
Db 693 CAGTACCGGATCCGGAACCTTTTGAAGCGCAACACCGGCAACGATGTGGCTGAGGCTCC 752
QY 841 ACCATCTGAGCTCTCAAGCGGGTGAAGAGTTGGCTGCAATCTTTCTCACTCAAGACAG 900
Db 753 ACCATCTGAGCTCTCAAGGAGGGTGAAGAGTCTGGTTCACTCACTGGAGACAG 812
QY 901 AACGGGCTCTTATGACCTTTACTGAGACAGAGCTCTTTTACGGGCTTCTTAATCTAT 960
Db 813 AATGAGCTCTTATCAACCTTTATTTGAGACCGAGCTGTTCACCGGCTTCTCAATCTAC 872
QY 961 GCGGACGAGATGACCCCAAGGATATGACATGCGCACGCG-CGGTCTTCAGGCAAGGA 1019
Db 873 GCTGATCAAGAGAACCCCAATGAGTATGACAAGCTGGGTTGAGCGTCAAGGCAAGGA 932
QY 1020 ACAAGCTTCTGAGCTTGGGCTTTCAGAGCAAGACCCCAACTGTGAGCTGGGGTGGGG 1079
Db 933 CTAAAGATTCGCAAGGGTGTGATAGAGAGAGTCTCTGAATGAGGCTGGGCACTGGCA 992
QY 1080 GGTGAGTGAAGGCTTCTAGCCTCAGGCTCAGCTCCGCGCTCTTTT 1129
Db 993 GTTCTTGGAGC--TTTATTCCAGGCAAGCTCTCTGTGCTGCTTT 1040

RESULT 5
US-09-552-204A-10
; Sequence 10, Application US/09552204A
; Patent No. 6620909
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; TITLE OF INVENTION: ADIPOCTE-SPECIFIC PROTEIN HOMOLOG ZACRP2
; FILE REFERENCE: 99-08
; CURRENT APPLICATION NUMBER: US/09/552,204A
; PRIORITY FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,207
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate nucleotide sequence encoding the
; NAME/KEY: variation
; LOCATION: (1)...(855)
```

```
OTHER INFORMATION: Each N is independently any nucleotide.
US-09-552-204A-10

Query Match 52.9%; Score 613.6; DB 4; Length 855;
Best Local Similarity 60.4%; Pred. No. 3,2e-148;
Matches 516; Conservative 170; Mismatches 168; Indels 0; Gaps 0;

QY 133 ATGATCCCTGGGCTGCTCTGAGCTGTGGCTTCCCTGCTGCTGAGTACCACTGCTTGGC 192
Db 1 ATGATTCCTTGGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 193 GCCTTGTCTGCGAGGACTTCCGAAAGGCTCCCTCAACTGTGCTGCAAGCTGCTGGC 252
Db 61 GCNTTTCGNNMGNMGNAYTTTGMNARAGNWSNCCNCAHYTGTNTGYMSNTNCCNGN 120
QY 253 CCCCAGGCGCCACCGCGCCCGCCCAAGAGGCGCCCAAGGCGCTCAGAGATGATGGGAGAAATG 312
Db 121 CCNAGAGGNCNCCNCGNCCNCCNCGNCCNCCNCGNCCNCGNCCNCGNCCNCGNCCNCGN 180
QY 313 GGCTTCTCTGCGAAAGAGCGCAAGATGAGCAAGAGCGGAGCCGCGGAGAG 372
Db 181 GNTTTCNAGNABARAYGNCARAYGNCAYGAGAYGNCAYGNCAYGNCAYGNCAYGNCAYG 240
QY 373 GAAAGTCACTTGGCGGCAAGGTAAACGGGGAAGCCAGAGCAAGAGGCAAGCCGGG 432
Db 241 GARGNCCNCCNCGNMGNCNCGNMGNCNCGNMGNCNCGNMGNCNCGNMGNCNCGNMGNCN 300
QY 433 GCCATTGGGGGGCTGGGCGCCCGCGGCGCCCAAGGGGTCAAGGTTACCCCGGAGCAT 492
Db 301 GCNATHTGNNMGNCCNCGNCCNCGNCCNCGNCCNCGNCCNCGNCCNCGNCCNCGNCCN 360
QY 493 GGCACACAGAGCAAGAGGAGGCGCCCAAGGCGCAAGAAAGGAGGCGAGCTCCAGGCGCC 552
Db 361 GGNACNCCNCGNABARABARGNCCNABARGNABARABARGNABARABARGNABARABAR 420
QY 553 TGCAGCTGTGGCAAGTGGCAATACCAAGTCTTCTCGGTGCACTGACCAAGACTAC 612
Db 421 TGYWNTGTGYWNSNGNCAYACNABRWSNCGNTTYSNGTNGCNGTNAACNABRWSNTAY 480
QY 613 CCAAGGAGCGGCGTCCCAAGTGTGAACAAGTCTGATGACAGAGGTGGCCACTAC 672
Db 481 CCNMGAGNMGNTNCCNATHTAATTTGAYAAATHTAATGAAVARGGNGNCAYTAY 540
QY 673 AATGCTTCCAGCGGCAAGTTCGTCTGCGGCGTGGCTGAGTCTACTACTTACCTACGAC 732
Db 541 AATGCMWSMWSNGNABARTTGTNTGYGNGTNCNGNABARTTAYATTTACNTAYGAY 600
QY 733 ATCAGCTGCGCAACAGCACTGGCCATGCGCTGTGTGCAACAGGCGCATGCCATC 792
Db 601 ATTAACNTYNGCNAAYAACAYTNGCATHGTGNTGTCAYAAVAGNCARTAYMGNAATH 660
QY 793 CGGACCTTTGAGCCACACCGGCAACCAAGATGAGGCTCAGGCTCAGCATCTGGGCT 852
Db 661 MGNACNTTTCAGGCAACCAAGATGAGGCTCAGGCTCAGGCTCAGCATCTGGGCT 720
QY 853 CTCAGCAGGAGTGAAGATTTGGCTGAGATCTTCTACTCAGAGCAAGAGGCTCTTTC 912
Db 721 YTTAARCAAGAGNAYARAGTNTGTYNCARATHTTTTAYTWSNABARABAYGNTNTTY 780
QY 913 TATGACCTTACTGACAGCAAGCTCTTTAGCGGCTTCCCTAATCTATGCGCACAGAT 972
Db 781 TAYGAYCMTAYTGACNAYMSNTNTTYACNGTNTTYTATHTAYGNCAYCARAY 840
QY 973 GACCCCAAGAGGT 986
Db 841 GAYCCNAAYGARGT 854

RESULT 6
US-09-188-930-19
; Sequence 19, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
```

```

?
? APPLICANT: Watson, James D.
? APPLICANT: Strachan, Lorna
? APPLICANT: Sleeman, Matthew
? APPLICANT: Onrust, Rene
? APPLICANT: Multison, James Greg
? TITLE OF INVENTION: Compositions Isolated From Skin Cells
? TITLE OF INVENTION: and Methods For Their Use
? FILE REFERENCE: 11000.101c1
? CURRENT APPLICATION NUMBER: US/09/188,930A
? CURRENT FILING DATE: 1998-11-09
? NUMBER OF SEQ. ID NOS: 348
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ. ID NO 19
? LENGTH: 750
? TYPE: DNA
? ORGANISM: Rat
US-09-188-930-19

```

Query Match	39.0%	Score 453;	DB 3;	Length 750;
Best Local Similarity	86.2%	Pred. No.	5.5e-107;	
Matches 501;	Conservative	0;	Mismatches 80;	Indels 0;
			Gaps	0

QY	121	GGGGTAAACCAACATATATCCCTGGGGGCTCTGGCCCTGGGCCCTCCCTGTGCTGAC	180
Db	170	AAGGTGACCAACATATATCTCTGGATGCTCTTGGCTTGAGCCCTTCCGTGTGCTGAC	229
QY	181	CCACTGCTTGGCGCCTTTGCTTCGACGGGACTTCGGAAGGCTCCCTCAACTGATCTGC	240
Db	230	CCAAATGCTTGGGCTTTGTCTTCGACGGGACTTCGAAAGGGTGTCTCAACTGATGTGC	289
QY	241	AGCTGCTTGGCCTCCCAAGGCCCAACCCGGCCCCCAGAGGCCCAAGGCCCTTAGGAATG	300
Db	290	AGTCTGCTGTGTCCCAAGGCCCAACTGTGCCTTCAGAGACCAAGATCTTCAGGAATG	349
QY	301	ATGGGACCAATAGGCGCTTTCCTTGGCAAAAGCGCCAAAGATGAGACAGACGCGACCCGAGG	360
Db	350	GTGGGAAAGATAGGCTTTCTCTGTAAAGATGAGCAAGCCAGACCGGACCGAAGCCAGGG	409
QY	361	GACAGCGGAGAGGAAGTTCACCTTGGCCGGAAGGTATACCGGAGAAAGCCAGGACCAAA	420
Db	410	GACAGTGAAGAAAGGTTCACCTTGGCAAGACAGGCAATCGAAGAAACAAAGACCAAA	469
QY	421	GGCAAAGCCGGGGCCATTGGGCGGGCTGGCCCCCTGTGCCCCCAAGGGGCTCAACGTTACC	480
Db	470	GGCAAAGCTGGGGCCATTGGGGAACCGGGTCTTCGAGAACCAAGGGGGCTCAATGTATACC	529
QY	481	CCCGGGAAGCATGAGCAACAGGCAAGAAAGGGGCCCAAGGGCAAGAAAGGAGGCCAAGC	540
Db	530	CCCGGGAAGCATGGTATACCGGGCAAGAAAGGGACCTTAAGGGCAAGAAAGGGAAACCTGGG	589
QY	541	CTCCGAGGCCCTGTGAGCTGTGGCAGTGGCCATACCAATGACGTTTCTCGGTGCGATG	600
Db	590	CTCCGAGGCCCTGTAGGTGTGGCGGCAGTAGCCGAGCCAAATCGGCCCTTTTTCGGTGGCGGTA	649
QY	601	ACCAAGAGCTACCCACGGGAGCGGCTGCCCATCAAGTTTGAACAAGTTCTGATGAACGAG	660
Db	650	ACCAAGAGTTTACCCAGTGAAGCGACTGCCCATCAAGTTTGAACAAGATTCTGATGAATGAG	709
QY	661	GGTGGCACTACAAATGCTTCCAGCGGCAAGTTGATGCGG	701
Db	710	GGAGGCACTACAAATGCACTCAAGTGGCAAGTTGTCTGCAAG	750

RESULT 7  
US-09-312-283C-19  
Sequence 19, Application US/09312283C  
Patent No. 6573095  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
Orntuf, Rene  
APPLICANT: Murrison, James G.

```

1  APPLICANT: Kimble, Krishanand D.
2  TITLE OF INVENTION: Compositions Isolated from Skin Cells
3  TITLE OF INVENTION: and Methods for Their Use
4  FILE REFERENCE: 11000-101162
5  CURRENT APPLICATION NUMBER: US/09/312,283C
6  CURRENT FILING DATE: 1999-05-14
7  NUMBER OF SEQ ID NOS: 425
8  SOFTWARE: FastSeq for Windows Version 4.0
9  SEQ ID NO 19
10 LENGTH: 750
11 TYPE: DNA
12 ORGANISM: Mouse
13 US-09-312-283C-19
14
15 Query Match          39.0%; Score 453; DB 4; Length
16 Best Local Similarity 86.2%; Pred. No. 5,5e-107;
17 Matches 501; Conservative 0; Mismatches 80; Indels

```

Query Match	39.0%	Score 453;	DB 4;	Length 750;
Best Local Similarity	86.2%	Pred. No. 5.5e-107;		
Matches 501; Conservative	0;	Mismatches 80;	Indels 0;	Gaps 0;

Oy	121	AGGGTAAACAACATATATCCCTGGAGTGCCTCTGGAGCTTGACCTCCCTGATGCTGTGAC	180
Db	170	AAAGTGACCAACATATATCTCTGGATGCTCTTGAGCTTGAGCCCTTCCGATGTGCTGTGAC	229
Oy	181	COACTGCTTGGACGCTTGTCTTGACAGGAACTTTCGCGAAAGCTCCCTCAACTGTGCTGC	240
Db	230	CCAAATGCTTGGTGCTTGTGCTTGACAGGAACTTTCGAAAGAGGTGTGCTCAACTGTGTGC	289
Oy	241	AGCTGTGCTTGGACCCCAAGAGGCCACACCGGACCCCAAGAGCCCAAGGCGCTCAAGAAAG	300
Db	290	AGTCTGCTTGTGTCCCAAGGCGCAACTGTGACCTTCAAGAGACACAGAAATCTTCAAGAAAG	349
Oy	301	ATGGGACAAATATGGGCTTCTCTTGCAAAAGACGCGCAAGATGACACACGACGCAACCGAGG	360
Db	350	GTGGGAAAGAAATGGGTTTCTCTGTAAAGATGACCAAGACGCGCAAGACCGAATCCAGAG	409
Oy	361	GACACGCGAGAGGAAGGTCCAATTGCGCGGACAGGTAACTCGGAGGAAAGCCACGACCAAAAG	420
Db	410	GACAGTGAAGAAAGGTCCAATTGCGCGGACAGGCAAGCAACCGAGGAAACCAAGACCAAAAG	469
Oy	421	GGCAAAAGCCGGGGCCATTGGGCGGGCTGCCCCGCTGCGCCCAAGGAGGGCTCAACGGTACC	480
Db	470	GGCAAAAGCTGGGGCCATTGGGAGACGGGCTCTTCAAGGACCCCAAGGAGGTCAAGTGTATCC	529
Oy	481	CCCGGGAAGCATGAGCACACAGGCAAGAGGGGCCCAAGGGCAAGAAAGGAGACAGAGC	540
Db	530	CCCGGGAACATGTGTATACCGGCGCAAGAGGSACTTAAGGGCAAGAAAGGGAACCTTGGG	589
Oy	541	CTCCAGGACCCCTGACAGCTGTGGCAGTGGCCATACCAAGTCAAGCTTCTCGGTGGCAGTG	600
Db	590	CTCCAGGACCCCTGTAGCTGTGGCGGACGTAAACGAGCCAAAGTCGAGCTTCTCGGTGGCGGTA	649
Oy	601	ACCAAGAGCTAACCAACGGGAGCGGCTGCCCATCAAGTTTGAACAAGTTCTGATGAACGAG	666
Db	650	ACCAAGAGTAAACCAACGTGAAGCGAATGCCCATCAAGTTTGAACAAGTTCTGATGAATGAG	709
Oy	661	GATGCGCACTAACATGCTTCCACGCGCAAGTTCGTCTGCGG	701
Db	710	GAAGGCACTAACATGCAATCCAGTGGCAAGTTCTGTCTGAG	750

```

; RESULT 8
; US-09-552-204A-11
; Sequence 11, Application US/09552204A
; Patent No. 6620909
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACRPT2
; FILE REFERENCE: 99-08
; CURRENT APPLICATION NUMBER: US/09/552,204A
; CURRENT FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,207
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 20

```

```
SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(360)
; NAME/KEY: variation
; LOCATION: (1)...(536)
; OTHER INFORMATION: Each N is independently any nucleotide.
US-09-552-204A-11
```

```
Query Match      24.9%; Score 289.6; DB 4; Length 536;
Best Local Similarity 76.5%; Pred. No. 4,7e-65;
Matches 404; Conservative 0; Mismatches 118; Indels 6; Gaps 4;
```

```
QY 631 ATCAAGTTTGACAAGATTCTGATGAAAGAGGGTGGCCCACTACATGCTTCCAGCGGCAAG 690
DB 1 ATCAAGTTTGACAAGATTCTGATGAAAGAGGGTGGCCCACTACATGCTTCCAGCGGCAAG 60
QY 691 TTGCTGGGGGGGCTGCTGGGATCTACT--ACTTCACTTACGATCAAGC--TGCGCAAC 747
DB 61 TTGCTGCAAGCGTGGCGGGGATCTAATTAACNTTACCTATGACATTACGCTTGGCCAC 120
QY 748 AAGCACTGGCCATCGGCTGTGCAACAAGCCAGTACCGCATCCGACCTTGTATGCC 807
DB 121 AAACACTGTCGATCGGCTGTGCAACAAGTGTCAAGTACCGCATCTTGTATGCC 180
QY 808 AACACCGGCAACCAAGATGTGGCTTCAACCAATCTGCTGTCAAGAGGGTAC 867
DB 181 AACACCGGCAACCAAGATGTGGCTTCAACCAATCTGCTGTCAAGAGGGTAC 240
QY 868 GAAGTTGGTGGAGATCTTCTAAGAGAGAGAGGGGCTTCTATGACCTTACTGG 927
DB 241 GAAGTTGGTGGAGATCTTCTAAGAGAGAGAGGGGCTTCTATGACCTTACTGG 300
QY 928 ACAGACAGCTCTTTACGGGCTTCTAATCTATGCGGACCAAGATGACCCCAAGAGTA 987
DB 301 ACCGACAGCTGTTCACCGGCTTCTCATCTACGTCGACCAAGAGAGAGCCCAAGAGTA 360
QY 988 TAGACATGCCACGG--CGGTCTCCAGGACGAGGAAACAAGCTTGTGACTGGCTTACAGA 1046
DB 361 TAGACATGCCACGG--CGGTCTCCAGGACGAGGAAACAAGCTTGTGACTGGCTTACAGA 418
QY 1047 GCAAGACCCCAACACTGTAGGCTGGGGGTGGAGTGAAGTGAAGCGGTTTACGCTCAGG 1106
DB 419 GGAAGATCCCTCGACTGGGCTGTGACATGACATCTTGGGATCTTTATTTCCAGGCAAG 478
QY 1107 CTCACCTCCCTCCGCTTTTTCCTTCCCTTCAATTAATCCAAACTTT 1154
DB 479 GCCTCCTCATTTGCTGCTTAAAAAAGAAATCAATTAATCCAAACTTT 526

RESULT 9
US-09-227-357-51
; Sequence 51, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
```

```
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 51
LENGTH: 1333
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (485)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (486)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
```



Db 643 ATGGG--NATHYAGCMNSNATHAARACGAYWSNACNTTYSNGNTTYTGTATAY 639  
QY 961 GCCGACGAGAT 972  
Db 700 WSNAGATYGGCAV 711

## RESULT 11

US-09-686-838B-10  
; Sequence 10, Application US/09686838B  
; Patent No. 6482612  
; GENERAL INFORMATION:  
; APPLICANT: Shepard, Paul O.  
; APPLICANT: Humes, Jacqueline M.  
; TITLE OF INVENTION: Adipocyte-Specific Protein Homologs  
; FILE REFERENCE: 97-49D1  
; CURRENT APPLICATION NUMBER: US/09/686,838B  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/140,804  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: US 60/056,983  
; PRIOR FILING DATE: 1997-08-26  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 729  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Degenerate nucleotide sequence encoding the zsig39  
; OTHER INFORMATION: polypeptide of SEQ ID NO:2.  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (1)..(729)  
; OTHER INFORMATION: n is any nucleotide  
US-09-686-838B-10

Query Match 8.7%; Score 101.2; DB 4; Length 729;  
Best Local Similarity 32.6%; Pred. No. 1.3e-16;

Matches 219; Conservative 88; Mismatches 350; Indels 15; Gaps 3;

QY 313 GCGTTTCCTGCGCAAGCGCCAGATGACACGACGCGACCGGGGGGACAGCGAGAG 372  
Db 43 GGMWNCNCNCTNYGAYGAAYAAATHCNWSNTYNTGCCNGNCAYCCNGGNYTN 102  
QY 373 GAAGTCCACTGCGCGGAGAGTAACCGGGGAAAGCAGACCAAGGCGAAAGCGGG 432  
Db 103 CCNGNACNCCNGNCAICAYAGGWSNCAAGNYTNCNGNMGNGAYGNGNGAYGN 162  
QY 433 GCCATTGGGCGGCTGGCCCCCTGGCCCCAGAGGGGTCAACGGTACCCCGGAGCAT 492  
Db 163 MNGAYGNGNCCNGNNGNCCNGNNGABABAGNABAGNNGNNGNCCNGNNTMCN 222  
QY 493 GGCAACACGAGCAAGAAAGGGGCCAAGGGCAAGAAAGGGAGCGAGGCTCCAGGCC 552  
Db 223 GGNCCMNGNAGNAYCCNGNCCMNGNAGNAGNCGNCGNCCNGNCCNACNGNCCN 282  
QY 553 TGAAGCTGTGGCAGTGGCCATCACTAGCTTTCTCGGTGGCAGTGAACAAGAG-- 608  
Db 283 GCGNGAGARTGTYSNNGNCCNCCMNGWSNCCNTTYSNCCNAAARMGWSNAGRWMSN 342  
QY 609 --CTAACCAAGGAGCGGCTGCCATCAAGTTTGAACAAGATTCTGATGAAGAGGGTGG 666  
Db 343 GTNCCNCCNCCMNSNAGNCCNYTTCCTTGTGATGNGTNTTNGNAAAGACAGAGN 402  
QY 667 CACTCAATGCTTCCAGCGGCAAGTTGCTGCGGCGTGCCTGGAGATCTACTTACCC 726  
Db 403 CATTAYGAGCNGTNAACNGNAAATYTAACGTGACGTGTCNGNNGTNTATYATYTCN 462  
QY 727 TAGGATCAACGCTGGCGCAACAGCACTGGCCATGGGCTGTGTGCAACAGCGCACTAC 786  
Db 463 GTNCAAGCNAACGNTATYMGNGCWSNNTYTCATTYGAATYTGTAARAAVGGAGWSN 522

QY 787 CGATCCGACCTTGTATGCAACACCGG-----CAACCAAGATGTGGCTCAGCTCC 840  
Db 523 ATHGCMWSNTTYTTCARTTYTYYGNGGNTGCGCAABACNGCMNSNTYMNAGNGN 582  
QY 841 ACCATCTGCTCTCAAGCAGGCTGACGAAGTTTGGCTGCAATCTTCTACTAGAGCAG 900  
Db 583 GCNATGTGNGNYTNGARCCNGARGAYCARGTGTGGTNCARGTNGNNGNGAYTAY 642  
QY 901 AACGGCTCTTCTATGACCTTACTGACAGACAGCTCTTACGGGCTTCTATCTAT 960  
Db 643 ATGGG--NATHYAGCMNSNATHAARACGAYWSNACNTTYSNGNTTYTGTATAY 639  
QY 961 GCCGACGAGAT 972  
Db 700 WSNAGATYGGCAV 711

## RESULT 12

US-09-140-804-9  
; Sequence 9, Application US/09140804  
; Patent No. 6197930  
; GENERAL INFORMATION:  
; APPLICANT: Shepard, Paul O.  
; APPLICANT: Humes, Jacqueline M.  
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS  
; FILE REFERENCE: 97-49  
; CURRENT APPLICATION NUMBER: US/09/140,804  
; CURRENT FILING DATE: 1998-08-26  
; EARLIER APPLICATION NUMBER: 60/056,983  
; EARLIER FILING DATE: 1997-08-26  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 4517  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-140-804-9

Query Match 8.3%; Score 96.8; DB 3; Length 4517;  
Best Local Similarity 49.7%; Pred. No. 3.5e-15;

Matches 319; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

QY 328 GAGCGCCAAAGTGAACGACGCGACCGGGGGGACAGCGAGAGGAGGTCACCTGCG 387  
Db 129 GCTTCGACAGTTGATGCGGGCATCCAGGGCATCCGGGCAATATGAGGGCCAGGCG 188  
QY 388 CGGACAGTAAACCGGGAAAGCCAGACCAAGAGCGGGCAATTTGGGCGGGCT 447  
Db 189 CGTATGGAAGATGACACCTGTGTGAAGAGGTGAAGAAAGATCCAGGTCTTATT 248  
QY 448 GGGCCCCGTCGCCCCCAAGGGGTCAACGGTACCCCGGGAAGATGACACAGGCAAG 507  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGAGATACCGGGCTGAAGTCCCGAGGCTTT 308  
QY 508 AAGGGGCCAAGGGCAAGAAAGGAGCAAGGCTCCAGGCCCTGCAAGCTGTGGCACT 567  
Db 309 CGGGAATCCAGGACGAGAAAGGAAAGCTGGAAGAGTGTCTATATATCCG-- 362  
QY 568 GAGCATTAACAGTCACTTCTCGGTGCGAGTGAACCAAGCTAACCAAGGAGCGGCTG 627  
Db 363 -----TCAGCATTCAGTGTGGATTTGGAATCTTACCTTATCCCAACATG 410  
QY 628 CCGATCAAGTTGAACAAGATTCTGATGAAGAGGCTGGCCTTCAAGCTTCCAGCGG 687  
Db 411 CCGATTCGCTTACCAAGATCTTCTAATCAAGAAACCACTATATAGCTTCACTGGT 470  
QY 688 AAGTGTCTGCGGCTGCTGGAGTCTACTTCACTTACAGACATCAAGCTGGCGAAC 747  
Db 471 AATTCACGTGCAAGATTCCTGGGCTGTACTTCTTACCAATCAAGCTATATAG 510  
QY 748 AAGCACTGGCATCGGCTGTGTGCAACAGGCAATCCAGTCCGACCTTGTATGCC 807



Db 531 AAGGATGTGAAGGTGACGCTCTTCAAGAGAGACAGGCTATGCTTCACTTATGATCAG 590  
Qy 808 AACACCGGCAACA---CGATGTGGCTTCAGGCTCCACCATCTGTGGCTTCAGAGGCT 864  
Db 591 TACGAGAAATATATGTGACACAGGCTCCGCTCTGTGCTTCGATCTGAGAGTGGC 650  
Qy 865 GACGAATTTGGCTGAGATCTTCTACTGAGAGAGACGGGCTCTTATGACCTTAC 924  
Db 651 GACCAAGTCTGGCTCCAGGCTATGAGAGAGAGAGGCTATGAGCTCTTATGCTATAT 710  
Qy 925 TGGACAGACAGCTCTTATGAGGCTTCTATCTATGCTATGCTATGCTATGCTATG 966  
Db 711 GACCAATGACTCCACCTTCAAGGCTTCTTCTTCACTATGCTATGCTATGCTATGCTATG 752

RESULT 13  
US-09-686-838B-9  
; Sequence 9, Application US/09686838B  
; Patent No. 6482612  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Humes, Jacqueline M.  
; TITLE OF INVENTION: Adipocyte-Specific Protein Homologs  
; FILE REFERENCE: 97-49D1  
; CURRENT APPLICATION NUMBER: US/09/686,838B  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/140,804  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: US 60/056,983  
; PRIOR FILING DATE: 1997-08-26  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PaeSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 4517  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-686-838B-9

Query Match 8.3%; Score 96.8; DB 4; Length 4517;  
Best Local Similarity 49.7%; Pred. No. 3.5e-15;  
Matches 319; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

Qy 328 GACGCGCAAGATGAGACAGACGCGGAGCCGGGGGAGACGAGAGAGAGAGGTCACCTGCGC 387  
Db 129 GCTTGCACAGGTTGATGAGCGGGGATCCAGGGCATCCGGCCATATAGGGGCCCGCAGGC 188  
Qy 388 CGGACAGGTTAACCGGGGAAAGCAGAGACAAAGGGCAAGCCGGGGCTATGGCGGGCT 447  
Db 189 CGGATGTGAGAGATGAGACCCCTGTGTGAGAGAGGTGAGAGAGATCCAGGCTTAT 248  
Qy 448 GGGCCCCGCGCCCAAGGGGGTCAACGTTACCCCGGAAAGCATGACACACAGGCAAG 507  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGAGTACCCGGGGCTGAGAGGTCCCGAGGCTTT 308  
Qy 508 AAGGGGCGCAAGGAG 567  
Db 309 CCGGAAATCCAAAGGAG 362  
Qy 568 GGGCATTACCAAGTCAAGCTTCTCGTGGAGTGAACCAAGAGCTACCCAGGGAGCGGCTG 627  
Db 363 -----TCAGATTCAAGTGTGGAGTTGAGAGCTTACGTTATCTTCCCAATG 410  
Qy 628 CCATCAAGTTTGAACAAGATTCTGATGAACAGAGGAGGAGTCAATGCTTCCAGGCGC 687  
Db 411 CCATTCGCTTTACAAAGATCTTCAATCAAGCAAAACATATGATGCTCCATGCT 470  
Qy 688 AAGTGTCTGCGCGGCTGAGATCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 747  
Db 471 AATTTCACTGCAACATTTCTGCGGCTGTACTTCTTCACTTCACTTCACTTCACTTCA 530  
Qy 748 AAGCACTGCGCAATCGGCTGTGCAACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 807  
Db 531 AAGGATGTGAAGGTGACGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 590

Qy 808 AACACCGGCAACA---CGATGTGGCTTCAGGCTCCACCATCTGTGGCTTCAGAGGCT 864  
Db 591 TACGAGAAATATATGTGACACAGGCTCCGCTCTGTGCTTCTGATCTGAGAGTGGC 650  
Qy 865 GACGAATTTGGCTGAGATCTTCTACTGAGAGAGACGGGCTCTTATGACCTTAC 924  
Db 651 GACCAAGTCTGGCTCCAGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710  
Qy 925 TGGACAGACAGCTCTTATGAGGCTTCTATCTATGCTATGCTATGCTATGCTATGCTATG 966  
Db 711 GACCAATGACTCCACCTTCAAGGCTTCTTCTTCACTATGCTATGCTATGCTATGCTATG 752

RESULT 14  
US-09-776-976-5  
; Sequence 5, Application US/09776976  
; Patent No. 6566332  
; GENERAL INFORMATION:  
; APPLICANT: Fruebis, Joachim  
; APPLICANT: Erickson, Mary Ruth  
; APPLICANT: Yen, Frances  
; APPLICANT: Bihain, Bernard  
; TITLE OF INVENTION: OB3 Globular Head and Uses Thereof for Decreasing Body Mass  
; FILE REFERENCE: 76,US,REG  
; CURRENT APPLICATION NUMBER: US/09/776,976  
; CURRENT FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 09/758,055  
; PRIOR FILING DATE: 2001-01-10  
; PRIOR APPLICATION NUMBER: US 60/176,228  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: US 60/198,087  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: US 60/299,881  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 4517  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-776-976-5

Query Match 8.3%; Score 96.8; DB 4; Length 4517;  
Best Local Similarity 49.7%; Pred. No. 3.5e-15;  
Matches 319; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

Qy 328 GACGCGCAAGATGAGACAGACGCGGAGCCGGGGGAGACGAGAGAGAGAGGTCACCTGCGC 387  
Db 129 GCTTGCACAGGTTGATGAGCGGGGATCCAGGGCATCCGGCCATATAGGGGCCCGCAGGC 188  
Qy 388 CGGACAGGTTAACCGGGGAAAGCAGAGACAAAGGGCAAGCCGGGGCTATGGCGGGCT 447  
Db 189 CGGATGTGAGAGATGAGACCCCTGTGTGAGAGAGGTGAGAGAGATCCAGGCTTAT 248  
Qy 448 GGGCCCCGCGCCCAAGGGGGTCAACGTTACCCCGGAAAGCATGACACACAGGCAAG 507  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGAGTACCCGGGGCTGAGAGGTCCCGAGGCTTT 308  
Qy 448 GGGCCCCGCGCCCAAGGGGGTCAACGTTACCCCGGAAAGCATGACACACAGGCAAG 507  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGAGTACCCGGGGCTGAGAGGTCCCGAGGCTTT 308  
Qy 508 AAGGGGCGCAAGGAG 567  
Db 309 CCGGAAATCCAAAGGAG 362  
Qy 568 GGGCATTACCAAGTCAAGCTTCTCGTGGAGTGAACCAAGAGCTACCCAGGGAGCGGCTG 627  
Db 363 -----TCAGATTCAAGTGTGGAGTTGAGAGCTTACGTTATCTTCCCAATG 410  
Qy 628 CCATCAAGTTTGAACAAGATTCTGATGAACAGAGGAGGAGTCAATGCTTCCAGGCGC 687  
Db 411 CCATTCGCTTTACAAAGATCTTCAATCAAGCAAAACATATGATGCTCCATGCT 470  
Qy 688 AAGTGTCTGCGCGGCTGAGATCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 747  
Db 531 AAGGATGTGAAGGTGACGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 590

Db 471 AATTCACCTGCAACATTCCTGGGCTGTACTTCTTATGCTACACATCAGATCTATATG 530  
Qy 748 AAGCACTGGGCATTCGGCTGTGTGCAACAGGCGAGTACCGCATCCGAGCTTTGATGCC 807  
Db 531 AAGGATGTGAAGTCACTCTTCAGAAAGACAGGCTATGCTTCCACTTATGATCAG 590  
Qy 808 AACACCGGCAACCA---CGATGTGGCTCAGGCTCCACATCCCTGCTCTCAAGAGGT 864  
Db 591 TACCAAGAAAATATGTGTGACACAGGCTCCGGGCTGTGTGCTCTGTGATCTGAGAGTGGC 650  
Qy 865 GACGAAGTTGGCTGTGAGATCTTCTACTCAGACAGAACGGGCTCTTCTATGACCTTAC 924  
Db 651 GACCAAGTGTGGCTCCAGGCTGTATGGGAAAGAGAGGCTATGAGACTTATGCTGATTAAT 710  
Qy 925 TGGAACAGACGCTCTTAAAGGCTTCTATATCTATGCCGAC 966  
Db 711 GACAAATGACTCCACTTCACAGGCTTCTTCTTACCATGAC 752

## RESULT 15

US-09-547-5  
; Sequence 5, Application US/0909547  
; Patent No. 6579852  
; GENERAL INFORMATION:  
; APPLICANT: Fruebis, Joachim  
; APPLICANT: Erickson, Mary Ruth  
; APPLICANT: Yen, Frances  
; APPLICANT: Bihain, Bernard  
; TITLE OF INVENTION: OB63 Globular Head and Uses Thereof for Decreasing Body Mass  
; FILE REFERENCE: 76.US6.CIP  
; CURRENT APPLICATION NUMBER: US/09/909,547  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: US 09/776,976  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 09/758,055  
; PRIOR FILING DATE: 2001-01-10  
; PRIOR APPLICATION NUMBER: US 60/299,881  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: US 60/198,087  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: US 60/176,228  
; PRIOR FILING DATE: 2000-01-14  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 4517  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-547-5

Query Match 8.3%; Score 96.8; DB 4; Length 4517;  
Best Local Similarity 49.7%; Pred. No. 3.5e-15;  
Matches 319; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

Qy 328 GACGGCCCAAGATGACACGAGCGGAGCCGGGGGAGACAGCGAGAGGAGTCCACTGGC 387  
Db 129 GCTTGACAGGTGTGATGGGGGATCCAGGGCATCCGGGCATATGGGGCCCCAGGC 188  
Qy 388 CGGACAGGTAAACGGGGAAAGCCAGAGACCAAGGGCAAAGCCGGGGCCATTGGGCGGGCT 447  
Db 189 CTGATGTGGAGAGATGGCACTCTGTGTGAGAAAGGTGAGAAAGAGATCCAGGCTTATTT 248  
Qy 448 GGGCCCCGTTGGCCCAAGGGGGTCAACGTTACCCCGGAAAGCATGGCACACAGGCAAG 507  
Db 249 GGTCTTAAGGAGACATCGGTGAACCGAGTACCCGGGGCTGAAGTCCCGAGGCTTT 308  
Qy 508 AAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGCTTCCAGGCCCTGACGCTGTGGCAGT 567  
Db 309 CCGGGAATCCAGGCAAGAAAGGAACTTGAAGAGTGTCTTATGTATACCGC----- 362  
Qy 568 GGGCATACCAAGTCAAGTCTTCTGCTGTGCACTGACCAAGAGCTAACCAAGGAGCGGCTG 627  
Db 363 -----TCAGCATTCAGTGTGGATTGGAAGCTTACGTTACTATCCCAACATG 410

Qy 628 CCCATCAAGTTTGACAAGATTTGTATGAACGAGGGTGGCCACTTACATAGTCTCCAGGCGC 687  
Db 411 CCCATTGCTTTACCAAGATCTTCTACATCAGAAACCACTATGATGCTCCACTGCT 470  
Qy 688 AAGTTGCTGCGGCGTGGCTGTGGATCTACTACTTCACTACGACATCAGCTGGCCAC 747  
Db 471 AATTCACCTGCAACATTCCTGGGCTGTACTACTTGTGCTTACCAATCAGATCTATATG 530  
Qy 748 AACACCTGGCCATTCGGCTGTGTGCAACAGGCAATCCGATCCGAGCTTTGATGCC 807  
Db 531 AAGGATGTGAAGTCAAGCTCTTCAAGAAAGACAGGCTATGCTTCACTTATGATCAG 590  
Qy 808 AACACCGGCAACCA---CGATGTGGCTCAGGCTCCACATCCCTGCTCTCAAGAGGT 864  
Db 591 TACCAAGAAAATATGTGTGACACAGGCTCCGGCTGTGTGCTCTGTGATCTGAGAGTGGC 650  
Qy 865 GACGAAGTTGGCTGTGAGATCTTCTACTCAGACAGAACGGGCTCTTCTATGACCTTAC 924  
Db 651 GACCAAGTGTGGCTCCAGGCTGTATGGGAAAGAGAGGCTATGAGACTTATGCTGATTAAT 710  
Qy 925 TGGAACAGACGCTCTTAAAGGCTTCTATATCTATGCCGAC 966  
Db 711 GACAAATGACTCCACTTCACAGGCTTCTTCTTACCATGAC 752

Search completed: December 25, 2004, 06:42:11  
Job time : 128 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2004, 05:29:28 ; Search time 708 Seconds  
(Without alignments)  
9130.890 Million cell updates/sec

Title: US-10-621-787-1

Perfect score: 1161  
Sequence: 1 ggaataactgctctggggcc.....aatccaactttattca 1161

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 4105333 seqs, 2784095677 residues

Total number of hits satisfying chosen parameters: 8210666

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/2/pubpna/PCY\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/2/pubpna/PCYUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 17: /cgn2\_6/prodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 18: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1161	100.0	1161	16	US-10-621-787-1
2	1154.2	99.4	1171	16	US-10-220-120-143
3	1148.4	98.9	1178	16	US-10-162-335-29
4	1120.8	96.5	1171	16	US-10-264-049-444
5	1084	93.4	1211	15	US-10-411-120-40
6	1082.8	93.3	1114	15	US-10-411-120-12
7	988.4	85.1	993	15	US-10-236-055A-5
8	703.6	60.6	885	15	US-10-236-055A-7
9	702.2	60.5	1107	10	US-09-866-050A-217
10	702.2	60.5	1107	14	US-10-152-661-217
11	701.6	60.4	1052	10	US-09-866-050A-358
12	701.6	60.4	1052	14	US-10-152-661-358

13	627.4	54.0	886	9	US-09-770-906-1	Sequence 1, Appl1
14	613.6	52.9	855	16	US-10-621-787-10	Sequence 10, Appl
15	507	43.7	855	9	US-09-770-906-3	Sequence 3, Appl1
16	453	39.0	750	10	US-09-866-050A-19	Sequence 19, Appl
17	453	39.0	750	14	US-10-152-661-19	Sequence 11, Appl
18	337.4	29.1	909	15	US-10-234-000-11	Sequence 11, Appl
19	309	26.6	1282	15	US-10-234-000-14	Sequence 14, Appl
20	295.4	25.4	870	15	US-10-203-708-7	Sequence 7, Appl1
21	295.4	25.4	870	16	US-10-451-168-105	Sequence 105, App
22	295.4	25.4	912	15	US-10-234-000-1	Sequence 1, Appl1
23	295.4	25.4	912	15	US-10-203-708-8	Sequence 8, Appl1
24	295.4	25.4	1297	18	US-09-728-952-85	Sequence 85, Appl
25	295.4	25.4	1297	18	US-10-758-846-58	Sequence 58, Appl
26	295.4	25.4	1336	15	US-10-411-120-11	Sequence 11, Appl
27	295.4	25.4	1387	18	US-10-758-846-57	Sequence 57, Appl
28	289.6	24.9	336	16	US-10-621-787-11	Sequence 11, Appl
29	253.2	21.8	1333	10	US-09-983-802-51	Sequence 51, Appl
30	253.2	21.8	1333	10	US-09-984-490-51	Sequence 51, Appl
31	253.2	21.8	1333	11	US-09-973-278-111	Sequence 11, App
32	253.2	21.8	1333	15	US-10-411-120-39	Sequence 39, Appl
33	195.2	16.8	266	16	US-10-242-535A-9251	Sequence 9251, Ap
34	195.2	16.8	266	16	US-10-085-783A-9251	Sequence 9251, Ap
35	185.4	16.0	252	16	US-10-242-535A-10450	Sequence 10450, A
36	185.4	16.0	252	16	US-10-085-783A-10450	Sequence 10450, A
37	174.4	15.0	894	16	US-10-423-584-3	Sequence 3, Appl1
38	173.2	14.9	999	16	US-10-423-582-3	Sequence 3, Appl1
39	160.4	13.8	277	16	US-10-242-535A-9290	Sequence 9290, Ap
40	160.4	13.8	277	16	US-10-085-783A-9290	Sequence 9290, Ap
41	133	11.5	1335	16	US-10-423-582-1	Sequence 1, Appl1
42	133	11.5	1530	16	US-10-411-120-49	Sequence 49, Appl
43	133	11.5	1530	18	US-10-726-699-41	Sequence 41, Appl
44	133	11.5	1545	15	US-10-411-120-21	Sequence 21, Appl
45	133	11.5	1545	18	US-10-726-699-18	Sequence 18, Appl

#### ALIGNMENTS

RESULT 1  
US-10-621-787-1  
; Sequence 1, Application US/10621787  
; Publication No. US20040024187A1  
; GENERAL INFORMATION:  
; APPLICANT: Piddington, Christopher S.  
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGY  
; FILE REFERENCE: 99-08D1  
; CURRENT APPLICATION NUMBER: US/10/621,787  
; CURRENT FILING DATE: 2003-07-17  
; PRIOR APPLICATION NUMBER: US 09/552,204  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: US 60/130,207  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1161  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (133)...(987)  
US-10-621-787-1

Query Match. 100.0%; Score 1161; DB 16; Length 1161;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATACTATGCTGGGGCGAGCGCTGCGCGGCTGCGCGTGAAGAAAGCGGAGAC 60  
Db 1 GGAATACTATGCTGGGGCGAGCGCTGCGCGGCTGCGCGTGAAGAAAGCGGAGAC 60

QY	61	GGGAGAGCCCGCGCAGAGCTTCCTTTGCTCCGGAGCGCCCTCGAGCGTGGCGGGCAGCGCG	120
Db	61	GGGAGAGCCCGCGCGAGAGCTTCTTGTGCTCCGGAGCGCCCTCGAGCGTGGCGGGCAGCGCG	120
QY	121	AGGGTACCACCATGATCCCTCTGGGTGCTCTGAGCTGTGCCTTCCCTGTGCTTGAC	180
Db	121	AGGGTACCACCATGATCCCTCTGGGTGCTCTGAGCTGTGCCTTCCCTGTGCTTGAC	180
QY	181	CCATGTGTTGGGCGCTTTTGTCTGCGAGGACCTTCCGAAAGGCTCCCTGAACTGGTCTGC	240
Db	181	CCATGTGTTGGGCGCTTTTGTCTGCGAGGACCTTCCGAAAGGCTCCCTGAACTGGTCTGC	240
QY	241	AGCTGCTGGGCGCCCGAGGGGCCAACCCGGGCCCCCAGAGAGCCCGAGGGCCCTCAGAAATG	300
Db	241	AGCTGCTGGGCGCCCGAGGGGCCAACCCGGGCCCCCAGAGAGCCCGAGGGCCCTCAGAAATG	300
QY	301	ATGGGACGATAGGAGCTTTTCTTGCCAAAGACCGGCGCAAGATGACACGACGCGACCCGGGG	360
Db	301	ATGGGACGATAGGAGCTTTTCTTGCCAAAGACCGGCGCAAGATGACACGACGCGACCCGGGG	360
QY	361	GACAGCGGAGAGAGAGTCACTTGCGCGGACAGGTAAACCGGGGAAAGCCAGAACAAAG	420
Db	361	GACAGCGGAGAGAGAGTCACTTGCGCGGACAGGTAAACCGGGGAAAGCCAGAACAAAG	420
QY	421	GGCAAAAGCCGGGGCCATTGGGGGGGCTGGCCCCCGTGGCCCCCAGAGGGGTCAAGGTACC	480
Db	421	GGCAAAAGCCGGGGCCATTGGGGGGGCTGGCCCCCGTGGCCCCCAGAGGGGTCAAGGTACC	480
QY	481	CCCGGGAACTATGGCACACAGGACAAAGAGGGGCCAAAGGGCCAAAGAGGGGAGCCAGGC	540
Db	481	CCCGGGAACTATGGCACACAGGACAAAGAGGGGCCAAAGGGCCAAAGAGGGGAGCCAGGC	540
QY	541	CTCCCAAGGCCCTGCAAGCTGTGGCAATGGCCAAATACAGCTTTCTCGGTGGCAATG	600
Db	541	CTCCCAAGGCCCTGCAAGCTGTGGCAATGGCCAAATACAGCTTTCTCGGTGGCAATG	600
QY	601	ACCAAGAGTATCCACACGGGAGCGGCTGCCCATCAAGTTTGAACAATTTCTAATGAACAG	660
Db	601	ACCAAGAGTATCCACACGGGAGCGGCTGCCCATCAAGTTTGAACAATTTCTAATGAACAG	660
QY	661	GGTGGCACTAACAATGCTTCCAGCGGACAAAGTTGCTGTGCGGGGCTGCGGANTCTATAC	720
Db	661	GGTGGCACTAACAATGCTTCCAGCGGACAAAGTTGCTGTGCGGGGCTGCGGANTCTATAC	720
QY	721	TTCACTTACGATCAAGCTGTGGCCAAACAAGCACTGGCCATGTGGCTGTGGTCAACAAGC	780
Db	721	TTCACTTACGATCAAGCTGTGGCCAAACAAGCACTGGCCATGTGGCTGTGGTCAACAAGC	780
QY	781	CAGTACCGGATCCGACCTTTTATGTCGCAACCCGGGCAACCAAGATGAGGCTCAGGCTCC	840
Db	781	CAGTACCGGATCCGACCTTTTATGTCGCAACCCGGGCAACCAAGATGAGGCTCAGGCTCC	840
QY	841	ACCATCTGAGCTTCAAGCAGGGGTGACGAATTTGGCTGACAGATTTTATCTCAGAGCAG	900
Db	841	ACCATCTGAGCTTCAAGCAGGGGTGACGAATTTGGCTGACAGATTTTATCTCAGAGCAG	900
QY	901	AACGGGCTTCTTATGACCTTTACTGACAGACAGCTTTTACGGGCTTCTTATCTAT	960
Db	901	AACGGGCTTCTTATGACCTTTACTGACAGACAGCTTTTACGGGCTTCTTATCTAT	960
QY	961	GCGGACACAGATGACCCCAACAAGATATGACATGCAACGGGGGTCTTCCAGGACGGGAA	1020
Db	961	GCGGACACAGATGACCCCAACAAGATATGACATGCAACGGGGGTCTTCCAGGACGGGAA	1020
QY	1021	CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGCCCAACAATGATGAGCTGGGGGGGGGG	1080
Db	1021	CAAGCTTCTGAGCTTGGGCTTACAGAGCAAGCCCAACAATGATGAGCTGGGGGGGGGG	1080
QY	1081	GTCGATGAGCGGTTCTAGCCTCAGGCTCACTCTCGGCTCTTTTTCCTCCCTTCAAT	1140
Db	1081	GTCGATGAGCGGTTCTAGCCTCAGGCTCACTCTCGGCTCTTTTTCCTCCCTTCAAT	1140
QY	1141	AAATCCAAACTTTTATTTCA	1161

```

DB          1141 AATCGAACCTTTTATTC 1161
|||||
RESULT 2
US-10-220-120-143
/ Sequence 143, Application US/10220120
/ Publication No. US20040048253A1
/ GENERAL INFORMATION:
/ APPLICATION: INCYTE GENOMICS, INC.
/ APPLICATION: PANZER, Scott R.
/ APPLICATION: SPIRO, Peter A.
/ APPLICATION: BANVILLE, Steven C.
/ APPLICATION: SHAH, Purvi
/ APPLICATION: CHALUP, Michael S.
/ APPLICATION: CHANG, Simon C.
/ APPLICATION: CHEN, Alice
/ APPLICATION: D'SA, Steven A.
/ APPLICATION: AMSHEY, Stefan
/ APPLICATION: DAHL, Christopher R.
/ APPLICATION: DAM, Tam C.
/ APPLICATION: DANIELS, Susan E.
/ APPLICATION: DUFOUR, Gerard E.
/ APPLICATION: FLORES, Vincent
/ APPLICATION: FONG, Willy T.
/ APPLICATION: GREENMALT, Lila B.
/ APPLICATION: HILLMAN, Jennifer L.
/ APPLICATION: JONES, Anissa L.
/ APPLICATION: LIU, Tommy F.
/ APPLICATION: ROSEBERRY, Ann M.
/ APPLICATION: ROSEN, Bruce H.
/ APPLICATION: RUSSO, Frank D.
/ APPLICATION: STOCKREHER, Theresa K.
/ APPLICATION: DAFFO, Abel
/ APPLICATION: WRIGHT, Rachel J.
/ APPLICATION: YAP, Pierre B.
/ APPLICATION: YU, Jimmy Y.
/ APPLICATION: BRADLEY, Diana L.
/ APPLICATION: BRATCHER, Shawn R.
/ APPLICATION: CHEN, Wensheng
/ APPLICATION: COHEN, Howard J.
/ APPLICATION: HODGSON, David M.
/ APPLICATION: LINCOLN, Stephen E.
/ APPLICATION: JACKSON, Stuart
/ TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: PT-1113 PCT
/ CURRENT APPLICATION NUMBER: US/10/220,120
/ CURRENT FILING DATE: 2002-08-26
/ PRIOR APPLICATION NUMBER: 60/184,777; 60/184,698; 60/184,770; 60/184,774;
/ 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
/ 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
/ 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
/ 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
/ 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
/ 60/205,324; 60/205,286
/ PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
/ 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
/ 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-05-17;
/ 2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;
/ 2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
/ 2000-05-17; 2000-05-17
/ NUMBER OF SEQ ID NOS: 422
/ SOFTWARE: PERL Program
/ SEQ ID NO 143
/ LENGTH: 1171
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:414307.1:2000FEB01
/ FEATURE:
/ NAME/KEY: unsure
/

```



LENGTH: 1178  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (4) .. (994)  
 US-10-162-335-29

Query Match 98.9%; Score 1148.4; DB 16; Length 1178;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1160; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

1 GGAATCTATGCTGGGGGCGAGCTGTGCCCCGGCTGCTCCGCTGAGAGAAAGCCGGAGC 60  
 Db 7 GGAATCTATGCTGGGGGCGAGCTGTGCCCCGGCTGCTCCGCTGAGAGAAAGCCGGAGC 66  
 QY 61 GCGGAGCCCGCGGAGAGCTTCTTGTCTCGGAGCCCTTGAACGTGGCGGAGCCGCG 120  
 Db 67 GCGGAGCCCGCGGAGAGCTTCTTGTCTCGGAGCCCTTGAACGTGGCGGAGCCGCG 126  
 QY 121 AGGGTAAACCACTGATCCCTGGGTGCTGCTGGCCCTGTCCTCCCTGCTGCTGCTGAC 180  
 Db 127 AGGGTAAACCACTGATCCCTGGGTGCTGCTGGCCCTGTCCTCCCTGCTGCTGCTGAC 186  
 QY 181 CCACTGCTGGCGCTTGTGCTCGAGGGAATTCCGAGAAAGCTCCCTCAACTGTGTGC 240  
 Db 187 CCACTGCTGGCGCTTGTGCTCGAGGGAATTCCGAGAAAGCTCCCTCAACTGTGTGC 246  
 QY 241 AGCTGCTGGCGCTTGTGCTCGAGGGAATTCCGAGAAAGCTCCCTCAACTGTGTGC 300  
 Db 247 AGCTGCTGGCGCTTGTGCTCGAGGGAATTCCGAGAAAGCTCCCTCAACTGTGTGC 306  
 QY 301 ATGGGAGCAATGGGCTTCTTGTGGAAGAAGCGGCAATGAGACAGACGCGAGCCGGGG 360  
 Db 307 ATGGGAGCAATGGGCTTCTTGTGGAAGAAGCGGCAATGAGACAGACGCGAGCCGGGG 366  
 QY 361 GACAGCGAGAGAGAGGTCCACTGTGCGGACAGGTAAACGGGGAAAGCCAGGACCAAG 420  
 Db 367 GACAGCGAGAGAGAGGTCCACTGTGCGGACAGGTAAACGGGGAAAGCCAGGACCAAG 426  
 QY 421 GCGAAGCCGGGACCATTTGGGCGGGCTGGCCCCCTGGTGGCCCAAGGGGTCAACGATAC 480  
 Db 427 GCGAAGCCGGGACCATTTGGGCGGGCTGGCCCCCTGGTGGCCCAAGGGGTCAACGATAC 486  
 QY 481 CCGGGAGAGATGGCAACAGAGCAAGAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGC 540  
 Db 487 CCGGGAGAGATGGCAACAGAGCAAGAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGC 546  
 QY 541 CTCCTCAGGCGCTGCAAGTGTGGCAATGACATCAAGTCAAGTCTTCTGCTGGCAGTG 600  
 Db 547 CTCCTCAGGCGCTTGTGAGTGTGGCAATGACATCAAGTCAAGTCTTCTGCTGGCAGTG 606  
 QY 601 ACCAAGCTTACCCAGGAGCGGCTGCCATCAAGTTTGAACAAGATTCTGATGAAGAG 660  
 Db 607 ACCAAGCTTACCCAGGAGCGGCTGCCATCAAGTTTGAACAAGATTCTGATGAAGAG 666  
 QY 661 GGTGGCACTAACATGCTTCCAGGCGGCAAGTGTCTGGCGGCGGCTGGAGATTATAC 720  
 Db 667 GGTGGCACTAACATGCTTCCAGGCGGCAAGTGTGTGGCGGCGGCTGGAGATTATAC 726  
 QY 721 TTCACTTACAGATCAAGCTGTGCAACAGCACTGTGCAATGGCTGTGTGCAACAGGC 780  
 Db 727 TTCACTTACAGATCAAGCTGTGCAACAGCACTGTGCAATGGCTGTGTGCAACAGGC 786  
 QY 781 CAGTACCGCATCCGGAATTTGATGCAACACCGGCAACAGATGTGGCTTCAAGGCTCC 840  
 Db 787 CAGTACCGCATCCGGAATTTGATGCAACACCGGCAACAGATGTGGCTTCAAGGCTCC 846  
 QY 841 ACCATCTGGCTCTCAAGAGGAGTGAAGAACTTGGTGTGCAAGATCTTCTACTGAGACAG 900  
 Db 847 ACCATCTGGCTCTCAAGAGGAGTGAAGAACTTGGTGTGCAAGATCTTCTACTGAGACAG 906  
 QY 901 AAGGGCTCTTATGACCTTACTGAGACAGACAGCTCTTTAGGGGCTTCTATATCTAT 960

Db 907 AAGGGCTCTTATGACCTTACTGAGACAGACAGCTCTTTAGGGGCTTCTATATCTAT 966  
 QY 961 GCCGACCAAGATATACCCCAACAGAGTATAGACATGCGACGCGGCTCTCCAGCAGGAA 1020  
 Db 967 GCCGACCAAGATATACCCCAACAGAGTATAGACATGCGACGCGGCTCTCCAGCAGGAA 1026  
 QY 1021 CAAGCTTGTGAATTTGGGCTTACAGAGCAAGACCCCAACTGTAGAGCTGGGGTGGAG 1080  
 Db 1027 CAAGCTTGTGAATTTGGGCTTACAGAGCAAGACCCCAACTGTAGAGCTGGGGTGGAG 1086  
 QY 1081 GTGAGTGAACGGTGTAGCTTACAGGCTCACTCTCCGCTC-TTTTTTCCTTTGAT 1139  
 Db 1087 GTGAGTGAACGGTGTAGCTTACAGGCTCACTCTCCGCTCTTTTTCCTTTGAT 1146  
 QY 1140 TAAATCAAACTTTTATTTCA 1161  
 Db 1147 TAAATCAAACTTTTATTTCA 1168

RESULT 4  
 US-10-264-049-444  
 ; Sequence 444, Application US/10264049  
 ; Publication No. US2004000579A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Birs et al.  
 ; TITLE OR INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P133P1  
 ; CURRENT APPLICATION NUMBER: US/10/264,049  
 ; CURRENT FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: PCT/US01/18569  
 ; PRIOR FILING DATE: 2001-06-07  
 ; PRIOR APPLICATION NUMBER: US 60/209,467  
 ; NUMBER OF SEQ ID NOS: 4360  
 ; SOFTWARE: PatentIn Ver. 3.1  
 ; SEQ ID NO 444  
 ; LENGTH: 1171  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-264-049-444

Query Match 96.5%; Score 1120.8; DB 16; Length 1171;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1137; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 18 GCCGAGCTCTGCTCCGCGCTGCTGCGCTGAGAAAGCCGGAGACGCC--CGCCGA 75  
 Db 1 GCCGAGCTCTGCTCCGCGCTGCTGCGCTGAGAAAGCCGGAGACGCCCGGACGAG 60  
 QY 76 GAGCTTCTTGTGCTCGGAGAGCCCTGAGCGTGGCGGAGCGCGAGAGGTAAACACATG 135  
 Db 61 GAGCTTCTTGTGCTCGGAGAGCCCTTGAAGTGGCGGAGCGCGAGAGGTAAACACATG 120  
 QY 136 ATCCCTGGGTGCTCTGGCTGTGCTGCTCCCTGTGCTGTGACCACTGCTTGGGCC 195  
 Db 121 ATCCCTGGGTGCTCTGGGTGTGCTGCTGCTCCCTGTGCTGTGACCACTGCTTGGGCC 180  
 QY 196 TTGCTGCGAGGAACTTCCGGAAGAGCTCCCTTCAACTGTGTGCTGAGGCTGCTGGCCC 255  
 Db 181 TTGCTGCGAGGAACTTCCGGAAGAGCTCCCTTCAACTGTGTGCTGAGGCTGCTGGCCC 240  
 QY 256 CAGGGCCACCCGGGCCCCCAGAGGCCCAAGAGCCCTCAGAGATGATGGAGCAATGGGC 315  
 Db 241 CAGGGCCACCCGGGCCCCCAGAGGCCCAAGAGCCCTCAGAGATGATGGAGCAATGGGC 300  
 QY 316 TTTCCTGGCAAGAGCGGCAAGATGAGACAGCGGACCGGGGGGACAGCGAGAGGAA 375  
 Db 301 TTTCCTGGCAAGAGCGGCAAGATGAGACAGCGGACCGGGGGGACAGCGAGAGGAA 360  
 QY 376 GGTCCACTGGCGGACAGGTAAACGGGGAAAGCCAGAACCAAGGGCAAGCGGGGCC 435  
 Db 361 GGTCCACTGGCGGACAGGTAAACGGGGAAAGCCAGAACCAAGGGCAAGCGGGGCC 420



```

QY 436 ATTGGGGGGGGTGGCCCCCGTGGCCCAAGGGGGTCAAGGTACCCCGGGAAAGCATGGC 495
Db 421 ATTGGGGGGGGTGGCCCCCGTGGCCCAAGGGGGTCAAGGTACCCCGGGAAAGCATGGC 480
QY 496 ACACAGAGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGCAGGCTCCAGGGCCCTGGC 555
Db 481 ACACAGAGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGCAGGCTCCAGGGCCCTGGC 540
QY 556 AGCTGTGGAGTGGCCATACCAAGTCACTTCTCGTGGCAGTACCAAGAGGTACCCA 615
Db 541 AGCTGTGGAGTGGCCATACCAAGTCACTTCTCGTGGCAGTACCAAGAGGTACCCA 600
QY 616 CGGAGAGGGCTGCCCATCAAGTTTGAACAAGATTCGATGAAAGGGGGTGGCCACTCAAT 675
Db 601 CGGAGAGGGCTGCCCATCAAGTTTGAACAAGATTCGATGAAAGGGGGTGGCCACTCAAT 660
QY 676 GCTTCCAGCGGCAAGTGGTGTGGGGTGGCTGGGAGTCTACTCACTTCAAGTACGAGATC 735
Db 661 GCTTCCAGCGGCAAGTGGTGTGGGGTGGCTGGGAGTCTACTCACTTCAAGTACGAGATC 720
QY 736 ACCTGTGGCAAGAGCACTGGCCATCGGCTGGTGAACAAGGCAAGTACCGGATCCGG 795
Db 721 ACCTGTGGCAAGAGCACTGGCCATCGGCTGGTGAACAAGGCAAGTACCGGATCCGG 780
QY 796 ACCTTTATGCAACACCGGCAACAGATGTGGCTCGAGGCTCAACATCTGGCTTC 855
Db 781 ACCTTTATGCAACACCGGCAACAGATGTGGCTCGAGGCTCAACATCTGGCTTC 840
QY 856 AAGCAGGGGTGCAAGATTGGCTGAGATCTTCTACTCAGAGCAAGAGGGCTCTTCTAT 915
Db 841 AAGCAGGGGTGCAAGATTGGCTGAGATCTTCTACTCAGAGCAAGAGGGCTCTTCTAT 900
QY 916 GACCTTACTGAGACAGACAGCTCTTTACAGGCTTCTTAATCTATGCGAGCAGATGAC 975
Db 901 GACCTTACTGAGACAGACAGCTCTTTACAGGCTTCTTAATCTATGCGAGCAGATGAC 960
QY 976 CCCAAGAGATATGACATGCGACAGGGGTCTCCAGGCAAGGAAACAAGCTTCTGACTT 1035
Db 961 CCCAAGAGATATGACATGCGACAGGGGTCTCCAGGCAAGGAAACAAGCTTCTGACTT 1020
QY 1036 GGGGTTACAGAGCAAGACCCCAACTGTAGCTGGGGGTGGGGGGTGGAGTGAAGGTT 1095
Db 1021 GGGGTTACAGAGCAAGACCCCAACTGTAGCTGGGGGTGGGGGGTGGAGTGAAGGTT 1080
QY 1096 CTAGGCTCAGGCTCACTCTCCGCTCTTTTCCCTTCAATTAATCCAAACTTTT 1155
Db 1081 CTAGGCTCAGGCTCACTCTCTGCTCTTTTCCCTTCAATTAATCCAAACTTTT 1140
QY 1156 TATTTCA 1161
Db 1141 TATTTCA 1146

```

```

; ORGANISM: Homo sapiens
US-10-411-120-40
Query Match 93.4%; Score 1084; DB 15; Length 1211;
Best Local Similarity 99.8%; Pred. No. 8e-310;
Matches 1084; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 76 GAGTTTCTTGTCTCCGAGACCCCTTGAAGTGGCGGCAACCCCGAGAGGTAAACCAATG 135
Db 1111 GAGTTTCTTGTCTCCGAGACCCCTTGAAGTGGCGGCAACCCCGAGAGGTAAACCAATG 1052
QY 136 ATCCCTGAGGCTCTGAGCTGAGCTTCCCTGAGTCTGTAAGCACTGTGGGCGC 195
Db 1051 ATCCCTGAGGCTCTGAGCTGAGCTTCCCTGAGTCTGTAAGCACTGTGGGCGC 992
QY 196 TTTGCTGCAAGGAATTCCGAAAAGGCTCCCTCAACTGTGTGCAAGCTTGGCTG 255
Db 991 TTTGCTGCAAGGAATTCCGAAAAGGCTCCCTCAACTGTGTGCAAGCTTGGCTG 932
QY 256 CAGGAGCCCAACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 315
Db 931 CAGGAGCCCAACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 872
QY 316 TTTCTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 375
Db 871 TTTCTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 812
QY 376 GGTTCACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 435
Db 811 GGTTCACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 752
QY 436 ATTGGGGGGGGTGGCCCCCGTGGCCCAAGGGGGTCAAGGTACCCCGGGAAAGCATGGC 495
Db 751 ATTGGGGGGGGTGGCCCCCGTGGCCCAAGGGGGTCAAGGTACCCCGGGAAAGCATGGC 692
QY 496 ACACAGAGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGCAGGCTCCAGGCGCTG 555
Db 691 ACACAGAGCAAGAAAGGGGGCCCAAGGGCAAGAAAGGGAGCAGGCTCCAGGCGCTG 632
QY 556 AGCTGTGGAGTGGCAATACCAAGTCACTTCTCGTGGCAAGTCAAGAGTCACTCA 615
Db 631 AGCTGTGGAGTGGCAATACCAAGTCACTTCTCGTGGCAAGTCAAGAGTCACTCA 572
QY 616 CGGAGAGGGCTGCCCATCAAGTTTGAACAAGATTCGATGAAAGGGGTGGCACTACAT 675
Db 571 CGGAGAGGGCTGCCCATCAAGTTTGAACAAGATTCGATGAAAGGGGTGGCACTACAT 512
QY 676 GCTTCCAGCGGCAAGTGGTGTGGGGTGGCTGGGAGTCTACTCACTTCAAGATC 735
Db 511 GCTTCCAGCGGCAAGTGGTGTGGGGTGGCTGGGAGTCTACTCACTTCAAGATC 452
QY 736 ACCTGTGGCAAGAGCACTGGCCATCGGCTGGTGAACAAGGCAAGTACCGGATCCG 795
Db 451 ACCTGTGGCAAGAGCACTGGCCATCGGCTGGTGAACAAGGCAAGTACCGGATCCG 392
QY 796 ACCTTTATGCAACACCGGCAACAGATGTGGCTCAAGGCTTCAACATCTGGCTTC 855
Db 391 ACCTTTATGCAACACCGGCAACAGATGTGGCTCAAGGCTTCAACATCTGGCTTC 332
QY 856 AAGCAGGGGTGCAAGATTGGCTGAGATCTTCACTCAGAGCAAGCGGCTTCTAT 915
Db 331 AAGCAGGGGTGCAAGATTGGCTGAGATCTTCACTCAGAGCAAGCGGCTTCTAT 272
QY 916 GACCTTACTGAGACAGACAGCTTCTTACAGGCTTCTTAATCTATGCGGCAAGAGTAC 975
Db 271 GACCTTACTGAGACAGACAGCTTCTTACAGGCTTCTTAATCTATGCGGCAAGAGTAC 212
QY 976 CCCAAGAGATATGACATGCGAGGGGTCTCCAGGCAAGGAAACAAGCTTCTGACTT 1035
Db 211 CCCAAGAGATATGACATGCGAGGGGTCTCCAGGCAAGGAAACAAGCTTCTGACTT 152
QY 1036 GGGGTTACAGAGCAAGACCCCAACTGTAGGCTGGGGGTGGAGTGAAGGTT 1095

```

```

RESULT 5
US-10-411-120-40/c
; Sequence 40, Application US/10411120
; Publication No. US20030224425A1
; GENERAL INFORMATION:
; APPLICANT: Blonder et al.
; TITLE OF INVENTION: ACRP-30-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P316P1
; CURRENT APPLICATION NUMBER: US/10/411.120
; PRIOR APPLICATION NUMBER: 2003-04-11
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: PCT/US02/32432
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 40
; LENGTH: 1211
; TYPE: DNA

```

Db	151	GGGCTTACAGACAGAACCCCAACAATGTAGGCTGGGGGGTGCAGTAGGCGGT	92
Qy	1096	CTAGCGTAGGGCTCAACCTGCTGGGCTCTTTTTTTCCTTCATTTAAATCCAAACCTTT	1155
Dd	91	CTAGCGTAGGGCTCACTCTCTGCGCTCTTTTTTCCCTTCATTAATCCAAACCTTT	32
Qy	1156	TATTTCATTTCA	1161
Dd	31	TATTTCATTTCA	26

```

RESULT 6
US-10-411-120-12
; Sequence 12, Application US/10411120
; Publication No. US20030224425A1
; GENERAL INFORMATION:
; APPLICANT: Blondel et al.
; TITLE OF INVENTION: ACRP-30-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT316P1
; CURRENT APPLICATION NUMBER: US/10/411,120
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/328,419
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: PCT/US02/32432
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 12
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-411-120-12

```

Query Match	93.3%	Score 1082.8	DB 15	Length 1114
Best Local Similarity	99.8%	Pred. No. 1.8e-309		
Matches 1084	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

QY	76	GAGCTTTCTTCTCCGGAAGCCCTCTGAGACGTGACGGGACGCGAGAGGTAAACAACATG	135
Db	9	GAGCTTCTTTTGTCTCCGGAAGCCCTCTGAGACGTGACGGGACGCGAGAGGTAAACAACATG	68
QY	136	ATCCCTTGAGGTGCTTCCTGAGCTGTGACCTTCCCTGTGTCTGTACCCAATGCTTGAGCGCC	195
Db	69	ATCCCTTGAGGTGCTTCCTGAGCTGTGACCTTCCCTGTGTCTGTACCCAATGCTTGAGCGCC	128
QY	196	TTTGCTGACAGGGAATTCCGGAAGAGCTCCCTCAACTGTGTGACGCTGCGTGGCGCC	255
Db	129	TTTGCTGACAGGGAATTCCGGAAGAGCTCCCTCAACTGTGTGACGCTGCGTGGCGCC	188
QY	256	CAGGACCCACCCGCGCCCTCCAGAGACCTCCAGGACCTTACGAAATGATGAGACGAATGAGC	315
Db	189	CAGGACCCACCCGCGCCCTCCAGAGACCTCCAGAGACCTTACGAAATGATGAGACGAATGAGC	248
QY	316	TTTCTGTGCAAAAGACGGCCAAAGATGGAACAGACGCGCGACCGGGGGGACAGCGGAAAGGA	375
Db	249	TTTCTGTGCAAAAGACGGCCAAAGATGGAACAGACGCGCGACCGGGGGGACAGCGGAAAGGA	308
QY	376	GGTCACTTGGCCCGGACAGGTAAACGGGGAAAGCCAGGACCAAAAGGCAAAAGCCGGAGCC	435
Db	309	GGTCACTTGGCCCGGACAGGTAAACGGGGAAAGCCAGGACCAAAAGGCAAAAGCCGGAGCC	368
QY	436	ATTGGGGCGGGCTGAGCCCTCGTGGCCCAAGGGGGGTCAACGTGATCCCTCGGGAAAGGATGGC	495
Db	369	ATTGGGGCGGGCTGAGCCCTCGTGGCCCAAGGGGGGTCAACGTGATCCCTCGGGAAAGGATGGC	428
QY	496	ACAACAGGCAAGAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGCTCCAGAGCCCTTGC	555
Db	429	ACAACAGGCAAGAGGGGCCCAAGGGCAAGAAAGGGAGCCAGGCTCCAGAGCCCTTGC	488
QY	556	AGCTGTGGCAATGGGCAATACCAATGCACTTTCTGTGTGTGACGTGACCAAGAGCTAACCA	615
Db	489	AGCTGTGGCAATGGGCAATACCAATGCACTTTCTGTGTGTGACGTGACCAAGAGCTAACCA	548

QY	616	CGGAGGGGGCGCCCATCAAGTTTGA	CAAGATTTGATGAA	CGAGGGGGGCACTACAT	675	
Db	549	CGGAGGGGGCGCCCATCAAGTTTGA	CAAGATTTGATGAA	CGAGGGGGGCACTACAT	608	
QY	676	GCTTCAGACGGCAAGTTCGTC	TGCGGGGTGCTTGAGATCTA	CTACTTCACCTACGACATC	735	
Db	609	GCTTCAGACGGCAAGTTCGTC	TGCGGGGTGCTTGAGATCTA	CTACTTCACCTACGACATC	668	
QY	736	ACGCTGGGCCAACAGCACCTG	GCATCGGCTGGTGCA	CAAGGCTAGTACCGCATCCGG	795	
Db	669	ACGCTGGGCCAACAGCACCTG	GCATCGGCTGGTGCA	CAAGGCTAGTACCGCATCCGG	728	
QY	796	ACCTTTGATGCAACACCGGCA	CCAGATGTGGCTCAGGCTCCA	CCATCTGGCTCTC	855	
Db	729	ACCTTTGATGCAACACCGGCA	CCAGATGTGGCTCAGGCTCCA	CCATCTGGCTCTC	788	
QY	856	AAGCAGGGGTGACGAAGTTTG	CTGCAGATCTTACTACAGACAA	CGGGCTCTTCAT	915	
Db	789	AAGCAGGGGTGACGAAGTTTG	CTGCAGATCTTACTACAGACAA	CGGGCTCTTCAT	848	
QY	916	GACCTTACTGGA	CAGACAGCCTCTTTACGGGCTT	CCTAATCTATGCGACCAAGATAC	975	
Db	849	GACCTTACTGGA	CAGACAGCCTCTTTACGGGCTT	CCTAATCTATGCGACCAAGATAC	908	
QY	976	CCCAACAGGATATGACA	TGCA	TGCGGGGTCTCCAGGCAAGGAA	CAAGCTTCGGAATT	1035
Db	909	CCCAACAGGATATGACA	TGCA	TGCGGGGTCTCCAGGCAAGGAA	CAAGCTTCGGAATT	968
QY	1036	GGGCTTACAGAGCAAGACCC	CAAACTGTAGGCTGGGGGTG	GGGGGGGTGAGTGA	CGGGTT	1095
Db	969	GGGCTTACAGAGCAAGACCC	CAAACTGTAGGCTGGGGGTG	GGGGGGGTGAGTGA	CGGGTT	1028
QY	1096	CTAGCCTCAGGCTCA	CTCCTCGGCTCTTTTTTCCC	CTTCATTAATCCAACTTTT	1155	
Db	1029	CTAGCCTCAGGCTCA	CTCCTCGGCTCTTTTTTCCC	CTTCATTAATCCAACTTTT	1088	
QY	1156	TATTTCA	1161			
Db	1089	TATTTCA	1094			

```

RESULT 7
US-10-236-055A-5
; Sequence 5, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christl
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236, 055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-055A-5

```

	Query Match	85.1%	Score 988.4	DB 15	Length 993
	Best Local Similarity	99.9%	Pred. No. 1.3e-281		
	Matches 989	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	1	GGAAACATGCTCGGAGCGCTCTGCGCTGCTGAGGAAAGCCGGGAC	60		
bb	4	GGAAACATGCTCGGAGCGCTCTGCGCGGCTGTGCGCTGAGGAAAGCCGGAC	63		

```

Qy 61 GCGGAGCCCCCGGAGAGCTTTTGTCTCCGAGGCCCTTGAGAGTGGCGGCGACCCGCG 120
Db 64 GCGGAGCCCCCGGAGAGCTTTTGTCTCCGAGGCCCTTGAGAGTGGCGGCGACCCGCG 123
Qy 121 AGGGTAACCAACATGATCCCTGGGATGCTCTGGGCTGTGCTCTCCCTGTGTCTGTAC 180
Db 124 AGGGTAACCAACATGATCCCTGGGATGCTCTGGGCTGTGCTCTCCCTGTGTGTCTGAC 183
Qy 181 CCATGCTTGGCGGCTTTTGTCTGAGAGGACTTCCGAAAAGGCTCCCTCAACTGCTGTC 240
Db 184 CCATGCTTGGCGGCTTTTGTCTGAGAGGACTTCCGAAAAGGCTCCCTCAACTGCTGTC 243
Qy 241 AGCTGCTTGGCGGCTTTTGTCTGAGAGGACTTCCGAAAAGGCTCCCTCAACTGCTGTC 300
Db 244 AGCTGCTTGGCGGCTTTTGTCTGAGAGGACTTCCGAAAAGGCTCCCTCAACTGCTGTC 303
Qy 301 ATGGGAGGAAATGGCTTTTCTGCGAAGAGCGGCAAGTGAACAAGCGGAGCCGAGG 360
Db 304 ATGGGAGGAAATGGCTTTTCTGCGAAGAGCGGCAAGTGAACAAGCGGAGCCGAGG 363
Qy 361 GACAGCGGAGGAGGAGGTCCAAGTGGCGGAGCAAGTAAACGGGAAAGCCAGAACAAAG 420
Db 364 GACAGCGGAGGAGGAGGTCCAAGTGGCGGAGCAAGTAAACGGGAAAGCCAGAACAAAG 423
Qy 421 GGCAGAACCGGGGCGCATTTGGGCGGGCTGGCCCGGTGGCCCAAGGGGGTCAAGGTACC 480
Db 424 GGCAGAACCGGGGCGCATTTGGGCGGGCTGGCCCGGTGGCCCAAGGGGGTCAAGGTACC 483
Qy 481 CCGGAGGAGGAGTGGCAACGAGGCAAGAGGGGGCCAGAGGCAAGAAAGGGAGCCAGGC 540
Db 484 CCGGAGGAGGAGTGGCAACGAGGCAAGAGGGGGCCAGAGGCAAGAAAGGGAGCCAGGC 543
Qy 541 CTCGAGGCGGCTTGCAGCTGTGGAGTGGCCATAAGTCAAGTCTTCTGCTGGGCAAGTG 600
Db 544 CTCGAGGCGGCTTGCAGCTGTGGAGTGGCCATAAGTCAAGTCTTCTGCTGGGCAAGTG 603
Qy 601 ACCAAGAGCTAACCAAGGAGGCGGCTGCCATCAAGTTTGAACAAGTTCTGATGAAGAG 660
Db 604 ACCAAGAGCTAACCAAGGAGGCGGCTGCCATCAAGTTTGAACAAGTTCTGATGAAGAG 663
Qy 661 GGATGCGCACTAACATGTCTTCAGAGGAGGAGTGGCTGGCGGCGGCTGGGATCTATCAC 720
Db 664 GGATGCGCACTAACATGTCTTCAGAGGAGGAGTGGCTGGCGGCGGCTGGGATCTATCAC 723
Qy 721 TTCACTTACGACATCAGCTGGCCCAACAGCACTGGCCATCGGCTGTGTGCAACAAGGC 780
Db 724 TTCACTTACGACATCAGCTGGCCCAACAGCACTGGCCATCGGCTGTGTGCAACAAGGC 783
Qy 781 CAGTACCGGATCCGGAACCTTTGATGCCAACACCGGCAACAGATGTGGCTCAGGCTCC 840
Db 784 CAGTACCGGATCCGGAACCTTTGATGCCAACACCGGCAACAGATGTGGCTCAGGCTCC 843
Qy 841 ACCATCTGGCTTCAAGAGGAGTGAAGAGTTTGGCTGAGATCTTCTACTCAGAGCAG 900
Db 844 ACCATCTGGCTTCAAGAGGAGTGAAGAGTTTGGCTGAGATCTTCTACTCAGAGCAG 903
Qy 901 AACGGGCTTCTTATGACCTTACTGACACAGCAAGCTCTTTAAGGGCTTCTATATCAT 960
Db 904 AACGGGCTTCTTATGACCTTACTGACACAGCAAGCTCTTTAAGGGCTTCTATATCAT 963
Qy 961 GCGGAGCAGATGACCCCAACAGAGTATAG 990
Db 964 GCGGAGCAGATGACCCCAACAGAGTATAG 993

```

```

RESULT 8
US-10-236-055A-7
; Sequence 7, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian

```

```

; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-236-055A-7

Query Match      60.6%; Score 703.6; DB 15; Length 885;
Beet Local Similarity 88.0%; Pred. No. 1.8e-197;
Matches 766; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 121 AGGGTAACCAACATGATCCCTGGGATGCTCTGGGCTGTGCTCTCCCTGTGTCTGTAC 180
Db 16 AGGGTAACCAACATGATCCCTGGGATGCTCTGGGCTGTGCTCTCCCTGTGTGTCTGAC 75
Qy 181 CCATGCTTGGCGGCTTTTGTCTGAGAGGACTTCCGAAAAGGCTCCCTCAACTGCTGTC 240
Db 76 CCATGCTTGGCGGCTTTTGTCTGAGAGGACTTCCGAAAAGGCGGGTCTCTCAACTGCTGTC 135
Qy 241 AGCTGCTTGGCGGCTTTTGTCTGAGAGGACTTCCGAAAAGGCTCCCTCAACTGCTGTC 300
Db 136 AGCTGCTTGGCTTCCGAGGCGGCTTCCGAGGAGCAAGATCTCTGAGAGTG 195
Qy 301 ATGGGAGGAAATGGCTTTTCTGCGAAGAGCGGCAAGTGAACAAGCGGAGCCGAGG 360
Db 196 ATGGGAGGAAATGGCTTTTCTGCGAAGAGCGGCAAGTGAACAAGCGGAGCCGAGG 255
Qy 361 GACAGCGGAGGAGGAGGTCCAAGTGGCGGAGCAAGTAAACGGGAAAGCCAGAACAAAG 420
Db 256 GACAGTGAAGAGGAGGAGGTCCAAGTGGCGGAGCAAGGCAACGCTGGAACAAAG 315
Qy 421 GGCAGAACCGGGGCGCATTTGGCGGCGGCTGGCCCGGCTGGCCCAAGGGGGTCAAGGTACC 480
Db 316 GGCAGAACCTGGGCGCATTTGGCGAGGCTGGCCCTGGAAGAACCAAGGGGGTCAAGGTACC 375
Qy 481 CCGGAGGAGGATGGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db 376 CCGGAGGAGGATGGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 435
Qy 541 CTCGAGGCGGCTTGCAGCTGTGGAGTGGCCATAAGTCAAGTCTTCTGCTGGGCAAGTG 600
Db 436 CTCGAGGCGGCTTGCAGCTGTGGAGTGGCCATAAGTCTTCTGCTGGGCAAGTG 495
Qy 601 ACCAAGAGCTAACCAAGGAGGCGGCTGCCATCAAGTTTGAACAAGTTCTGATGAAGAG 660
Db 496 ACCAAGAGCTAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 555
Qy 661 GGATGCGCACTAACATGTCTTCAGAGGAGGAGTGGCTGGCGGCTGTGAGATCTATATCAT 720
Db 556 GGATGCGCACTAACATGTCTTCAGAGGAGGAGGAGTGGCTGGCGGCTGTGAGATCTATATCAT 615
Qy 721 TTCACTTACGACATCAGCTGGCCCAACAGCACTGGCCATCGGCTGTGTGCAACAAGGC 780
Db 616 TTCACTTACGACATCAGCTGGCCCAACAGCACTGGCCATCGGCTGTGTGCAACAAGGC 675
Qy 781 CAGTACCGGATCCGGAACCTTTGATGCCAACACCGGCAACAGATGTGGCTCAGGCTCC 840
Db 676 CAGTACCGGATCCGGAACCTTTGATGCCAACACCGGCAACAGATGTGGCTCAGGCTCC 735
Qy 841 ACCATCTGGCTTCAAGAGGAGTGAAGAGTGTGGCTGAGATCTTCTACTCAGAGCAG 900
Db 736 ACCATCTGGCTTCAAGAGGAGTGAAGAGTGTGGCTGAGATCTTCTACTCAGAGCAG 795

```

QY 901 AACGGCTCTTCTATGACCTTACTGACAGACAGCTCTTACGGGCTTCTATCTAT 960  
DB 796 AATGGCTCTTCTATGACCTTACTGACAGACAGCTCTTACCGGCTTCTATCTAC 855  
QY 961 GCCGACCAAGATGACCCCAACGAGATATAG 990  
DB 856 GCTGACCAAGGAGACCCCAACGAGATATAG 885

RESULT 9  
US-09-866-050A-217  
; Sequence 217, Application US/09866050A  
; Publication No. US20030040471A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Muriison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OR INVENTION: Compositions Isolated From Skin Cells  
; TITLE OR INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c40  
; CURRENT APPLICATION NUMBER: US/09/866.050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 217  
; LENGTH: 1107  
; TYPE: DNA  
; ORGANISM: Rat  
US-09-866-050A-217

Query Match 60.5%; Score 702.2; DB 10; Length 1107;  
Best Local Similarity 84.7%; Pred. No. 5e-197;  
Matches 800; Conservative 0; Mismatches 143; Indels 2; Gaps 1;

QY 121 AGGTAACACACATGATCCCTGGGATGCTCTGAGCTGTGCTCCCTGCTGCTGAC 180  
DB 162 AAGGTGACACACATGATCTCTGATGCTCTGAGCTGTGCTCCCTGCTGCTGAC 221  
QY 181 CCATGCTGTGAGCTTGTGCTGACAGGACCTTCGGAAAGGCTCCCTCAACTGTGCTGC 240  
DB 222 CCATGCTGTGAGCTTGTGCTGACAGGACCTTCGGAAAGGCTTCCTCAACTGTGCTGC 281  
QY 241 AGCTGCTGTGAGCTTGTGCTGACAGGACCTTCGGAAAGGCTTCCTCAACTGTGCTGC 300  
DB 282 AGTGTGCTGTGCTTGTGCTGACAGGACCTTCGGAAAGGCTTCCTCAACTGTGCTGC 341  
QY 301 ATGGGACGAATGGCTTCTGTCGAAAGACGCAAGATGACAGACGGGACCGGGGG 360  
DB 342 GTGGGAAAGATGGCTTCTGTCGAAAGACGCAAGATGACAGACGGGACCGGGGG 401  
QY 361 GACACGCGAAGAAAGGTCCACTGCGCGAAGATGACAGACGGGACCGGGGG 420  
DB 402 GACAGTGAAGAAAGGTCCACTGCGCGAAGATGACAGACGGGACCGGGGG 461  
QY 421 GGCAAAACCGGGGCAATTTGGGCGGGCTGCGCCCTGCGCCCAAGGGGCTACAGGTACC 480  
DB 462 GGCAAAACCGGGGCAATTTGGGCGGGCTGCGCCCTGCGCCCAAGGGGCTACAGGTACC 521  
QY 481 CCGGGAAGCATGACACAGCAAGAAAGGGGCCAAGGSCAAGAAAGGGGACCGAGGC 540  
DB 522 CCGGGAAGCATGACACAGCAAGAAAGGGGCCAAGGSCAAGAAAGGGGACCGAGGC 581  
QY 541 CTCCGAGCCCTTCACTGTCGAGTGGCAATGCAAGTCAAGTCTTCTGCTGCGAGTG 600  
DB 582 CTCCGAGCCCTTCACTGTCGAGTGGCAATGCAAGTCAAGTCTTCTGCTGCGAGTG 641  
QY 601 ACAGAGAGCTTACCAAGGAGAGGCTGCGCCATCAAGTTTGAACAAGTTTGAACAAG 660  
DB 642 ACAGAGAGCTTACCAAGGAGAGGCTGCGCCATCAAGTTTGAACAAGTTTGAACAAG 701

QY 661 GGTGGCACTACATGCTTCCAGCGGCAAGTTGCTGCGGCGGTGCTGGGATCTACTAC 720  
DB 702 GAGGGCACTACATGCTTCCAGCGGCAAGTTGCTGCGGCGGTGCTGGGATCTACTAC 761  
QY 721 TTCACTTACATCAAGCTGTGGCCCAACAGCACTGTGCATGTGGCTGTGTGACAGGC 780  
DB 762 TTCACTTACATCAAGCTGTGGCCCAACAGCACTGTGCATGTGGCTGTGTGACAGGC 821  
QY 781 CAGTACCGCATCCGAGCTTGTGATGCAACACCGGCAACCAAGATGTGGCTGTGCTCC 840  
DB 822 CAGTACCGCATCCGAGCTTGTGATGCAACACCGGCAACCAAGATGTGGCTGTGCTCC 881  
QY 841 ACATCTGCTCTCAAGCAGGAGTGAAGATTTGCTGCAAGTCTTCTACTAGACAG 900  
DB 882 ACATCTGCTCTCAAGCAGGAGTGAAGATTTGCTGCAAGTCTTCTACTAGACAG 941  
QY 901 AACGGCTCTTCTATGACCTTACTGACAGACCTCTTTACGGGCTTCTATCTAT 960  
DB 942 AATGACTCTTCTATGACCTTACTGACAGACCTTCTTACCGGCTTCTATCTAT 1001  
QY 961 GCCGACCAAGATGACCCCAACGAGTATGACATGCCAGG--CGTCTCCAGGCGAGG 1018  
DB 1002 GCTGATCAAGAGACCCCAATGAGTATGACAGCAAGCTGGGGTGTGAGCGCTCCAGG 1061  
QY 1019 AACAGCTTCTGACTTGGGCTTACAGAGCAAGACCCCAACTG 1063  
DB 1062 ACTAGATTCGCAAGGCTGTGATGAGAGATCTGTGACTG 1106

RESULT 10  
US-10-152-661-217  
; Sequence 217, Application US/10152661  
; Publication No. US20030022835A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Muriison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OR INVENTION: Compositions Isolated From Skin Cells  
; TITLE OR INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c5  
; CURRENT APPLICATION NUMBER: US/10/152.661  
; CURRENT FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: 09/866.050  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/221.232  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: 60/206.650  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 09/312.283  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051  
; PRIOR FILING DATE: 1999-04-29  
; PRIOR APPLICATION NUMBER: 09/188.930  
; PRIOR FILING DATE: 1998-11-09  
; PRIOR APPLICATION NUMBER: 09/069.726  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 217  
; LENGTH: 1107  
; TYPE: DNA  
; ORGANISM: Rat  
US-10-152-661-217

Query Match 60.5%; Score 702.2; DB 14; Length 1107;  
Best Local Similarity 84.7%; Pred. No. 5e-197;  
Matches 800; Conservative 0; Mismatches 143; Indels 2; Gaps 1;

QY 121 AGGTAACACACATGATCCCTGGGATGCTCTGAGCTGTGCTCCCTGCTGCTGAC 180

```

Db 162 AAGGTGACACACATGATCTCTGATGCTCTTGAGCTCTGCTCTCTGCTGCTGAC 221
Qy 181 CCACTGCTTGCGCGCTTTGCTGCGAGGACTTCGGAAGGCTCCCTCAACTGCTGCG 240
Db 222 CCAATGCTTGCTGCTGCTGCTGCGAGGACTTCGGAAGGCTGCTCAACTGCTGCG 281
Qy 241 AGCTGCTGCTGCTGCTGCTGCGAGGACTTCGGAAGGCTGCTCAACTGCTGCG 300
Db 282 AGCTGCTGCTGCTGCTGCGAGGACTTCGGAAGGCTGCTCAACTGCTGCG 341
Qy 301 ATGGGACGAATGGGCTTTCTGCGCAAGAAGGCTGCGCAAGTGAACAGCGGAGCCGAGG 360
Db 342 GTGGGAAGAAATGGGCTTTCTGCGCAAGAAGGCTGCGCAAGTGAACAGCGGAGCCGAGG 401
Qy 361 GACAGCGGAGAGAGAGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGG 420
Db 402 GACAGTGAAGAGAGAGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGG 461
Qy 421 GCGAAGGCGGCGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGG 480
Db 462 GCGAAGGCGGCGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGG 521
Qy 481 CCGGGAAGAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 540
Db 522 CCGGGAAGAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 581
Qy 541 CTCGCAAGGCGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGG 600
Db 582 CTCGCAAGGCGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGG 641
Qy 601 ACCAAGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 660
Db 642 ACCAAGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 701
Qy 661 GGTGCGCAATGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 720
Db 702 GGTGCGCAATGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 761
Qy 721 TTACCTGACATGACATGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGG 780
Db 762 TTACCTGACATGACATGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGG 821
Qy 781 CAGTACCGCATGCGGAGCTTTGATGCGCAACCGGAGCCGAGGAGCCGAGGAGCCGAGG 840
Db 822 CAGTACCGCATGCGGAGCTTTGATGCGCAACCGGAGCCGAGGAGCCGAGGAGCCGAGG 881
Qy 841 ACCATCTGCTGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 900
Db 882 ACCATCTGCTGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 941
Qy 901 AACGGGCTCTTATGACCTTATGACAGGAGCTTATGACAGGAGCTTATGACAGGAG 960
Db 942 AATGGAATCTTATGACCTTATGACAGGAGCTTATGACAGGAGCTTATGACAGGAG 1001
Qy 961 GCGGACGAGATGACCTTATGACAGGAGCTTATGACAGGAGCTTATGACAGGAG 1018
Db 1002 GCGGACGAGATGACCTTATGACAGGAGCTTATGACAGGAGCTTATGACAGGAG 1061
Qy 1019 AACGAGCTTATGAGCTTATGACAGGAGCTTATGACAGGAGCTTATGACAGGAG 1063
Db 1062 ACTGAGATCTGCGCAAGGAGCTTATGACAGGAGCTTATGACAGGAGCTTATGAG 1106

```

## RESULT 11

```

; Sequence 358, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Jorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene

```

```

; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ. ID NOS: 725
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 358
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Rat
; US-09-866-050A-358

```

```

Query Match 60.4%; Score 701.6; DB 10; Length 1052;
Best Local Similarity 82.9%; Pred. No. 7,4e-197;
Matches 837; Conservative 0; Mismatches 169; Indels 4; Gaps 3;

```

```

Qy 121 AAGGTACACACATGATCTCTGATGCTCTTGAGCTCTGCTCTCTGCTGCTGAC 180
Db 34 AAGGTACACACATGATCTCTGATGCTCTTGAGCTCTGCTCTCTGCTGCTGAC 93
Qy 181 CCACTGCTTGCGCGCTTTGCTGCGAGGACTTCGGAAGGCTCCCTCAACTGCTGCG 240
Db 94 CCAATGCTTGCTGCTGCTGCTGCGAGGACTTCGGAAGGCTGCTCAACTGCTGCG 153
Qy 241 AGCTGCTGCTGCTGCTGCGAGGACTTCGGAAGGCTGCTCAACTGCTGCG 300
Db 154 AGCTGCTGCTGCTGCTGCGAGGACTTCGGAAGGCTGCTCAACTGCTGCG 212
Qy 301 ATGGGACGAATGGGCTTTCTGCGCAAGAAGGCTGCGCAAGTGAACAGCGGAGCCGAGG 360
Db 213 ATGGGACGAATGGGCTTTCTGCGCAAGAAGGCTGCGCAAGTGAACAGCGGAGCCGAGG 272
Qy 361 GACAGCGGAGAGAGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 420
Db 273 GACAGTGAAGAGAGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 332
Qy 421 GCGAAGGCGGCGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 480
Db 333 GCGAAGGCGGCGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 392
Qy 481 CCGGGAAGAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGG 540
Db 393 CCGGGAAGAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGG 452
Qy 541 CTCGCAAGGCGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 600
Db 453 CTCGCAAGGCGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 512
Qy 601 ACCAAGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGG 660
Db 513 ACCAAGGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGG 572
Qy 661 GGTGCGCAATGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGG 720
Db 573 GGTGCGCAATGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGG 632
Qy 721 TTACCTGACATGACATGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 780
Db 633 TTACCTGACATGACATGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGG 692
Qy 781 CAGTACCGCATGCGGAGCTTTGATGCGCAACCGGAGCCGAGGAGCCGAGGAGCCGAGG 840
Db 693 CAGTACCGCATGCGGAGCTTTGATGCGCAACCGGAGCCGAGGAGCCGAGGAGCCGAGG 752
Qy 841 ACCATCTGCTGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGG 900
Db 753 ACCATCTGCTGCTGCGCAAGTGAACAGCGGAGCCGAGGAGCCGAGGAGCCGAGGAGCCGAGG 812
Qy 901 AACGGGCTCTTATGACCTTATGACAGGAGCTTATGACAGGAGCTTATGACAGGAGCTTATGAG 960

```









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:08:11 ; Search time 73.0538 Seconds  
(without alignments) 375.364 Million cell updates/sec

Title: US-10-621-787-2

Perfect score: 1582

Sequence: 1 MIPWVLLACALPCADPLG.....TDSLFTGFLIYADDDNEV 285

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 100 summaries

Database :

1: PIR.79:\*  
2: p1r1:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	516.5	32.6	680	1 CGHUIB	collagen alpha 1(X
2	511	32.3	680	2 S31216	collagen alpha 1(X
3	502.5	31.8	674	2 S31301	collagen alpha 1(X
4	502	31.7	674	2 S23297	collagen alpha 1(X
5	443.5	28.0	635	2 A57131	collagen alpha 2(V
6	443.5	28.0	743	1 S23779	collagen alpha 1(V
7	439	27.7	744	1 A34246	collagen alpha 1(V
8	435	27.5	744	1 S23298	collagen alpha 1(V
9	431.5	27.3	744	2 UC4708	gelatin-binding 28
10	431	27.2	744	2 S15435	collagen alpha 1(V
11	420	26.5	423	2 A55797	collagen alpha 1(V
12	398.5	25.2	219	2 T14782	hypothetical prote
13	393	24.8	246	2 S29328	complement subcomp
14	376.5	23.8	253	1 C1H0B8	complement subcomp
15	369	23.3	245	2 S19018	complement subcomp
16	369	23.3	253	2 I49560	complement C1q B c
17	366	23.1	245	1 C1H0C	complement subcomp
18	357	22.6	253	2 S49158	complement protein
19	348.5	22.0	245	1 C1H0A	complement subcomp
20	311.5	19.7	1464	1 CGHUIS	collagen alpha 1(I
21	310	19.6	1453	2 S21626	collagen alpha 1(I
22	309	19.5	671	1 CGRTIS	collagen alpha 1(I
23	306	19.3	215	2 B48150	hibernation-relate
24	306	19.3	1042	1 CGCHIS	collagen alpha 1(I
25	303.5	19.2	779	1 CGB01S	collagen alpha 1(I
26	300.5	19.0	299	2 T24833	hypothetical prote
27	296.5	18.7	266	2 T22706	hypothetical prote
28	296.5	18.7	299	2 T22705	hypothetical prote
29	296.5	18.7	1690	1 CGHUIB	collagen alpha 4(I

30	294	18.6	1843	2 S18803	collagen alpha 1(V
31	293.5	18.6	1049	1 B4097S	collagen alpha 1(I
32	290.5	18.4	633	2 B40983	collagen alpha 1(X
33	289.5	18.3	281	2 T32765	hypothetical prote
34	289.5	18.3	1492	2 A40333	collagen alpha 1(I
35	288.5	18.2	1024	2 S18251	collagen alpha 1(X
36	288.5	18.2	1486	1 B40333	collagen alpha 1(I
37	288	18.2	1838	1 CGHUIV	collagen alpha 1(V
38	287.5	18.2	300	2 T24482	hypothetical prote
39	287.5	18.2	319	2 T32250	hypothetical prote
40	287.5	18.2	920	2 B34493	collagen alpha 1(I
41	287.5	18.2	2944	2 A54849	collagen alpha 1(V
42	286.5	18.1	1466	1 CGHUTL	collagen alpha 1(I
43	286	18.1	1487	1 CGHUC6	collagen alpha 1(I
44	285.5	18.0	140	2 A05249	collagen alpha 1(I
45	285.5	18.0	299	2 T29956	hypothetical prote
46	285.5	18.0	1464	2 S59856	collagen alpha 1(I
47	285	18.0	298	2 UC1448	collagen col-34 -
48	285	18.0	886	2 I50694	collagen alpha 1(I
49	285	18.0	1418	2 T45467	collagen alpha 1(I
50	285	18.0	1691	1 S22937	collagen alpha 5(I
51	283.5	17.9	310	2 I50696	collagen alpha 1(I
52	283	17.9	1419	2 A41182	collagen alpha 1(I
53	283	17.9	1487	2 B41182	collagen alpha 1(I
54	282	17.8	295	2 T22833	hypothetical prote
55	282	17.8	673	1 CGB06C	collagen alpha 1(I
56	282	17.8	888	2 S28791	collagen alpha 1(X
57	282	17.8	1142	2 UX0369	collagen alpha 1(X
58	282	17.8	1806	1 CGHUIE	collagen alpha 1(X
59	281.5	17.8	688	2 A53330	collagen alpha 2(I
60	281	17.8	453	2 S18804	collagen alpha 4(I
61	281	17.8	964	1 CGCH2S	collagen alpha 2(I
62	281	17.8	1388	2 A53317	collagen alpha 1(X
63	280.5	17.7	283	2 T19141	hypothetical prote
64	280.5	17.7	615	2 A05269	collagen alpha 1(I
65	280.5	17.7	636	2 S41067	collagen alpha 1(I
66	280	17.7	1747	2 A45974	collagen alpha 1(X
67	279.5	17.7	921	2 S40495	collagen alpha 1(I
68	278.5	17.6	201	2 T46404	hypothetical prote
69	278.5	17.6	1028	1 CGHUIA	collagen alpha 1(V
70	278	17.6	1366	1 CGHUIE	collagen alpha 2(I
71	277.5	17.5	324	2 T18763	hypothetical prote
72	277.5	17.5	488	2 A27353	collagen alpha 1(I
73	277.5	17.5	920	2 A45748	collagen alpha 1(V
74	277.5	17.5	1707	2 A35326	collagen alpha 2(I
75	277	17.5	296	2 T21070	hypothetical prote
76	277	17.5	1603	2 S23810	collagen alpha 1(X
77	276.5	17.5	286	2 S34665	collagen, cuticula
78	276.5	17.5	299	2 T25407	hypothetical prote
79	276	17.4	1019	1 A32856	collagen alpha 1(V
80	276	17.4	1373	1 A43281	collagen alpha 2(I
81	276	17.4	1414	1 S23809	collagen alpha 2(I
82	275.5	17.4	302	2 A31921	collagen dpv-13 pr
83	275	17.4	1497	2 I49607	procollagen type V
84	275	17.4	1758	2 T29350	hypothetical prote
85	274.5	17.4	623	2 A45137	collagen alpha 4(I
86	274.5	17.4	730	2 A36226	collagen alpha 1 c
87	274.5	17.4	921	2 S42617	collagen alpha 1(I
88	274.5	17.4	1669	1 CGMS4B	collagen alpha 1(I
89	274	17.3	324	2 T33149	hypothetical prote
90	273.5	17.3	324	2 A31920	collagen src-1 pre
91	273.5	17.3	438	2 S53787	collagen alpha cha
92	273.5	17.3	1546	1 CGHUIB	collagen alpha 2(X
93	273	17.3	283	2 US0170	collagen col-19 -
94	273	17.3	289	2 T27708	hypothetical prote
95	273	17.3	290	2 A32249	collagen - sea urc
96	273	17.3	341	2 T16296	hypothetical prote
97	273	17.3	366	2 S11449	collagen short cha
98	273	17.3	918	2 S23377	collagen alpha 2(V
99	273	17.3	1022	2 S04111	collagen alpha 2(V
100	273	17.3	1670	1 CGHUIB	collagen alpha 3(I

## ALIGNMENTS

## RESULT 1

CGR010  
collagen alpha 1(X) chain precursor - human  
N/Alternate names: procollagen alpha 1(X) chain  
C/Species: Homo sapiens (man)  
C/Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text change 09-Jul-2004  
C/Accession: S26396; S30086; S15826; S18249; A43901; I51870; S21856  
R/Reichenberger, B.; Beler, F.; Luvallie, P.; Olsen, B.R.; von der Mark, K.; Bertling, W.  
FEBS Lett. 311, 305-310, 1992  
A/Title: Genomic organization and full-length cDNA sequence of human collagen X.  
A/Reference number: S26396; MUID:93012005; PMID:1397333  
A/Accession: S26396  
A/Molecule type: DNA  
A/Residues: 1-680 <RE1>  
A/Cross-references: UNIPROT:Q03692; EMBL:X68952; EMBL:X72578; EMBL:X72579; EMBL:X72580; R/Apte, S.S.  
submitted to the EMBL Data Library, March 1992  
A/Reference number: S30085  
A/Accession: S30085  
A/Molecule type: DNA  
A/Residues: 'TTPYGVWCWCLL', 52-680 <APT>  
A/Cross-references: EMBL:X65120; NID:923129  
A/Note: The initial difference is probably due to translation of an intronic sequence  
R/Apte, S.; Mettel, M.G.; Olsen, B.R.  
FEBS Lett. 282, 393-396, 1991  
A/Title: Cloning of human alpha-1(X) collagen DNA and localization of the COL10A1 gene  
A/Reference number: S15826; MUID:91243838; PMID:2037056  
A/Accession: S15826  
A/Molecule type: DNA  
A/Residues: 561-647 'G', 649-666 <AP2>  
A/Cross-references: EMBL:X56879; NID:930013; PIDN:CAA41686.1; PID:930014  
R/Thomas, J.T.; Cresswell, C.J.; Raehn, B.; Nicolai, H.; Jones, T.; Solomon, E.; Grant, M.  
Biochem. J. 280, 617-623, 1991  
A/Title: The human collagen X gene. Complete primary translated sequence and chromosomal  
A/Reference number: S18249; MUID:92109659; PMID:1764025  
A/Accession: S18249  
A/Molecule type: DNA  
A/Residues: 1-26, 'T', 28-680 <THO>  
A/Cross-references: EMBL:X60382; NID:930094; PIDN:CAA2933.1; PID:930095  
A/Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 532-ALA  
R/Reichenberger, E.; Aigner, T.; von der Mark, K.; Stoss, H.; Bertling, W.  
Dev. Biol. 148, 562-572, 1991  
A/Title: In situ hybridization studies on the expression of type X collagen in fetal human  
A/Reference number: A43901; MUID:92077285; PMID:1743401  
A/Accession: A43901  
A/Molecule type: mRNA  
A/Residues: 547-656 <RE2>  
A/Cross-references: GB:M74050; GB:D57494; NID:9339884; PIDN:AAA61221.1; PID:9553796  
A/Note: sequence extracted from NCBI backbone (NCBI:69012, NCBI:69014)  
R/Mallis, G.A.; Raehn, B.; Sweetman, W.A.; Thomas, J.T.; Super, M.; Evans, G.; Grant, M.E.  
Am. J. Hum. Genet. 54, 169-178, 1994  
A/Title: Amino acid substitutions of conserved residues in the carboxyl-terminal domain  
pe Schmid.  
A/Reference number: I51870; MUID:94136476; PMID:8304336  
A/Accession: I51870  
A/Status: translated from GB/EMBL/DBDJB  
A/Molecule type: mRNA  
A/Residues: 520-597, 'D', 599-680 <NAL>  
A/Cross-references: GB:S68531; NID:9545180; PIDN:AAC60615.1; PID:9545181  
A/Note: mutant sequence from patient with metaphyseal chondrodysplasia type Schmid  
A/Note: a second mutant sequence with 614-Pro is also described  
C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
ed and subsequently O-glycosylated.  
C/Genetics:  
A/Gene: GDB:COL10A1  
A/Cross-references: GDB:128635; OMIM:120110  
A/Map position: 6q21-6q22  
A/Intons: 52/1  
A/Note: a defect in this gene may cause Schmid metaphyseal chondrodysplasia  
C/Complex: type X collagen may be a homotrimer

## C/Function:

A/Description: structural component of extracellular fibrous polymer specifically and tre  
be important for skeletogenesis  
C/Superfamily: collagen alpha 1(VIII) chain, complement C1q carboxyl-terminal homology  
C/Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyllysine; t  
F/1-18/Domain: signal sequence #status predicted <SIG>  
F/19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>  
F/19-56/Domain: amino-terminal nonhelical #status predicted <NC2>  
F/52-519/Region: interrupted helical  
F/520-680/Domain: amino-terminal nonhelical #status predicted <NC1>  
F/553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>  
F/617/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match 32.6%; Score 516.5; DB 1; Length 680;  
Best Local Similarity 41.5%; Pred. No. 5,8e-30;  
Matches 120; Conservative 29; Mismatches 93; Indels 47; Gaps 8;

Qy 26 DFRKSPQVLCSLPGPGPP--GPGAPGSGMGMRPGKDGQDGDHGDGDSGEG 82  
Db 397 DGRKGNP---GLPGKGDGPGVGGPPGLPGVPAGAKGPHGNHGEAGPRGAPGIPGTRG 452  
Qy 83 PPRGTGKRGKRGKRGKAGAIGRAGPRG--PKGVNTPGKRGTPGKKGKKGEPGLRGP 140  
Db 453 PIGPPGIGPPGSGKGPSPGPGAGIATKGLNGPT--GPGPPGRHSGBPGLGPG 509  
Qy 141 -----CSCSGHT---KSAFSAVATKSPYRRLPT 167  
Db 510 PGPSPGPGQAVMEGRFKAGGRPSLSGTPLVASNGVTPGVPVAFYILSKAPALGTP 569  
Qy 168 KPDKLINMEGSHYNASSGKFCVCGPIYFTYDITLANKKLAIGLVHNGQYRIRTPDANT 227  
Db 570 PFDKILVNRQOHYDPTGTFPTCOIPGIYYSYVHVHVGTHVWGLYKNGTPWVYTDVET 629  
Qy 228 -GNHDAVSSGTLALKQGDVWLCIIFYSEONGHFDYPTWDSLFETGLI 275  
Db 630 KGLVDASGSALITDLTENDQWMLQPVASNGLYSSRY-VHSSFSGFLV 677

## RESULT 2

S31216  
collagen alpha 1(X) chain precursor - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text change 09-Jul-2004  
C/Accession: S31216; S28807; S22215; S30127; I48299; S26397; S31830  
R/Kong, R.Y.C.; Kwan, K.M.; Lau, E.T.; Thomas, J.T.; Boot-Handford, R.P.; Grant, M.E.; Cf  
Eur. J. Biochem. 213, 99-111, 1993  
A/Title: Intron-exon structure, alternative use of promoter and expression of the mouse c  
A/Reference number: S31216; MUID:93238750; PMID:8477738  
A/Accession: S31216  
A/Molecule type: DNA  
A/Residues: 1-680 <KON>  
A/Cross-references: UNIPROT:Q05306; EMBL:Z21610; NID:949793; PIDN:CAA79736.1; PID:949794  
R/Bilma, K.; Berola, I.; Rosati, R.; Metcraeranta, M.; Garofalo, S.; Peraelae, M.; de Cron  
Biochem. J. 289, 247-253, 1993  
A/Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and exp  
A/Reference number: S28807; MUID:93143676; PMID:8424763  
A/Accession: S28807  
A/Molecule type: DNA  
A/Residues: 1-285, 'A', 287-680 <ELI>  
A/Cross-references: EMBL:X67348; NID:950480; PIDN:CAA47763.1; PID:950481  
R/Bilma, K.; Metcraeranta, M.; Kallio, J.; Peraelae, M.; Berola, I.; Garofalo, S.; de Cron  
Biochim. Biophys. Acta 1130, 78-80, 1992  
A/Title: Specific hybridization probes for mouse alpha-2(I)X and alpha-1(X) collagen mRNA  
A/Reference number: S22215; MUID:92182017; PMID:1543751  
A/Accession: S22215  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 385-450, 'K', 452-627 <ELA>  
A/Cross-references: EMBL:X63013; NID:949795; PIDN:CAA44741.1; PID:949796  
R/Apte, S.S.; Olsen, B.R.  
Matrix 13, 165-179, 1993  
A/Title: Characterization of the mouse type X collagen gene.  
A/Reference number: S30127; MUID:93261348; PMID:8492743



F:453,456/Modified site: hydroxyproline (Pro) #status experimental  
F:611/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.7%; Score 502; DB 2; Length 674;

Best Local Similarity 39.8%; Pred. No. 6,4e-29;

Matches 113; Conservative 29; Mismatches 93; Indels 49; Gaps 6;

```

Qy 38 LPPGQGP-----GPPAGPSPGMMGMPGPGKKGODGHODRSDSEBPG 85
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 389 LPPGKGPGLBPGQGDTHGAPPLPGLPBPVPGQKGPVINGEPGPPSPGIRPFG 448
Qy 86 RTGNRGPSPKKAIGRAPRPGKVGNGTPGKHGTPGKKKGPXKKGPEGLPGR----- 140
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 449 PPGMPGAPGAKGAGAPGLPBPAG-IATKGLRGMPPGPPGPPKNGSGEPGLPGRPPPG 507
Qy 141 -----CSCGSGHTK-----SAPSAVNTSYSPREPLPIKFDK 171
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 508 PPGSTPPEGVYKESRELSGMSFPMKAGANQALTMPPVSAFTVLSXAPGATVPKIFDK 567
Qy 172 ILMEGGHYNASGKPGVPGIYFTYDITLANKHLAIGLVHNGQYRIRTFDA-NTGNH 230
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 568 ILVNRQGHDPRTIFTCRIPLGLYFSYHAKGTWVWVALLYKNGSPVMTYIDYQKYL 627
Qy 231 DVASGSTITALKQGDVWLTQIFYSEONGLFYDPYMTDSLFTGFL 274
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 628 DQASGSAVIDLMENDQVWLQLPNSESNGLYSSEY-VHSSFSGFL 670

```

#### RESULT 5

A57131 collagen alpha 2(VIII) chain - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 03-Nov-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004

C:Accession: A57131

R:Muragaki, Y.; Jaccenko, O.; Abte, S.; Mattei, M.G.; Ninomiya, Y.; Olsen, B.R.

J. Biol. Chem. 266, 7721-7727, 1991

A:Title: The alpha2(VIII) collagen gene. A novel member of the short chain collagen faml

A:Reference number: A57131; MUID:91210292; PMID:2019595

A:Accession: A57131

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-635 <MUR>

C:Cross-references: UNIPROT:P25067; GB:M60832; NID:9177178; PIDN:AAA62822.1; PID:9177179

C:Genetics:

A:Gene: GDB:COL8A2

A:Cross-references: GDB:127812; OMIM:120252

A:Map position: 1p34.3-1p32.3

C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

F:1-11/Domain: amino-terminal nonhelical (fragment) #status predicted <NC2>

F:12-468/Region: interrupted helical

F:469-635/Domain: carboxyl-terminal nonhelical #status predicted <NC7>

F:508-634/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 28.0%; Score 443.5; DB 2; Length 635;

Best Local Similarity 26.5%; Pred. No. 1e-24;

Matches 113; Conservative 33; Mismatches 100; Indels 181; Gaps 7;

```

Qy 29 KGSQQL-----VCSLPGPGP-----PQPPAPPSG 55
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 207 KGPBGVGVGVAAGLPQPGPSGAKGEPGTRPGLIGFTGYGMFGLPQPKDRPAG 266
Qy 56 MMGMPGPPKXQGDGHGDRSDSEBPGPGRTPGRKGPKKAIGRAPRPGKGV-- 113
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 267 VPGILGRGREGELDESGQPGSLGPPGLPSAGLPGRGRPPGLAGBAGPGGPPVPG 326
Qy 114 -----NGTPGKHGTPGKKGPXK----- 131
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 327 IRPDQPSGLAGKPGVGERGLPAHAGPPGTPGKSGEGTGRPGPGVAGALGQKDLG 386
Qy 132 -----KGPGLPGPSCG----- 144
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 387 LPQPGILRGPSPGLGLQGPAPGIPGQGLPGLKGPBGPLPBPBGGRABPGTAGRGPBGV 446

```

Qy 145 -----SGHTK 149

Db 447 PGSPGITGPBGLGPBPAGPABETGIALHLNPGVGEAGVADKGGKPPGGLGSLNAT 506

Qy 150 SAPSAVATKSYPRERLPIKFDKILMEGGHYNASGKPFVCGVPGIYFTYDITLANKHLA 209

Db 507 PAFPAVLITSLPLPASGMVKFEDRLYNGHSGYNPATGIFTCPVGGVYFAHVHVKTNW 566

Qy 210 IGVHNGQYRIRTFDA-NTGNHVASGSTITALKQGDVWLTQIFYSEONGLFYDPYMTDS 268

Db 567 VALTKNNVPATYTYIDEXKKYLDQASGAVLQLRPNQVWVQLPSDANGL-YSTETIHS 625

Qy 269 LFTGFLI 275

Db 626 SFGSGL 632

#### RESULT 6

S23779

collagen alpha 1(VIII) chain - mouse

C:Species: Mus musculus (house mouse)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: S23779

R:Muragaki, Y.; Shiofa, C.; Inoue, M.; Ooshima, A.; Olsen, B.R.; Ninomiya, Y.

Eur. J. Biochem. 207, 895-902, 1992

A:Title: Alpha-1(VIII)-collagen gene transcripts encode a short-chain collagen polypeptic

A:Reference number: S23779; MUID:92362626; PMID:1499564

A:Accession: S23779

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-743 <MUR>

C:Cross-references: UNIPROT:Q00780; EMBL:X66976; NID:950493; PIDN:CAA47387.1; PID:913599;

C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

F:616-742/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 28.0%; Score 443.5; DB 1; Length 743;

Best Local Similarity 31.9%; Pred. No. 1.2e-24;

Matches 103; Conservative 36; Mismatches 97; Indels 87; Gaps 6;

```

Qy 40 GPGPPPPAPAPSPSGMMGMGR-----PQKQGDHDDDRGDSGEPPGRT 87
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 420 GPGPPPPKPSBPLQGPGRPLGVEVGPGRMGLPRLPGKSGKHGLPGLPVGGLL 479
Qy 88 GNRGKP-----GPKGKAIGRAPRPGKVGNGTPGK----- 119
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 480 GPKGEPITPDQGLQGPPLGIVGSPGRIGPPGICPPKBPGLPBPBPVGKPGVAG 539
Qy 120 -HGTPGKKGPKKGBEPGLPSCS----- 142
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 540 LHPGPKRGLGPGQGLPLGPPGPPGPAPVMPPTSPQGEVLPDMGLGIDGVKTPHA 599
Qy 143 -----CGSGHTKAFSAVTKSYPRELPIKFDKILMEGGHYNASGKPFVCGVPGI 194
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 600 YAGKKGHGGPAVEMPEFTLTLTPPPVGAIVKFDLLNQHONTNPPGTGIFTCVBPV 659
Qy 195 YFTYDITLANKHLAIGLVHNGQYRIRTFDA-NTGNHVDVSGSTITALKQGDVWLTQIFY 253
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 660 YFAIYHCHKGAVWVALLFKNNRPMWYTYIDYKKGFLDQASGAVLLLRGDDVFLQNP 719
Qy 254 SEONGLEFYDPYMTDSLFTGFLIY 276
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 720 EQAAGLYAGQY-VHSSFSGYLLY 741

```

#### RESULT 7

A34246

collagen alpha 1(VIII) chain precursor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: A34246

R:Yamaguchi, N.; Benya, P.D.; van der Raast, M.; Ninomiya, Y.

J. Biol. Chem. 264, 16022-16029, 1989

A:Title: The cloning and sequencing of alpha1(VIII) collagen cDNAs demonstrate that type



omains similar to those of type X collagen.

A:Reference number: A34246; MUID:89380199; PMID:2476437

A:Accession: A34246

A:Molecule type: mRNA

A:Residues: 1-744 <YAM>

A:Cross-references: UNIPROT:P14282; GB:J05042; NID:G164895; PIDN:AAA11204.1; PID:G164896

C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-744/Product: signal sequence #status predicted <MAT>

F:21-744/Region: amino-terminal nonhelical

F:118-571/Region: interrupted helical

F:572-744/Region: carboxyl-terminal nonhelical

F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 27.7%; Score 439; DB 1; Length 744;

Best Local Similarity 31.8%; Pred. No. 2.6e-24;

Matches 113; Conservative 33; Mismatches 99; Indels 110; Gaps 9;

30 GSPOLVCSLPGPQP-----GPPGAPSPSGMGMGPPGKD---GDDGH 71

390 GEPGLP-GIPEPMPPGAIIGPPKSGGIVGPPGPPGKBPGLQGPFGKGLGEGVP 448

72 DGDGDSGSEBGPRTGNRKP-----GPKGK-----AGA 101

449 PGIRGLPPIGPKBAGHKGRLPGVPGLLGPKGEPPIPDQGLQGPPIGRTGPGGP 508

102 IGRGPPGPKGVNGTPPK-----KTPGKPKGKKGKGGPGLP----- 140

509 IGPPIGPKBEPGLPBPFPVGVKPGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query 102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140

509 IGPPIGPKBPGLPBPFPFGVKGKGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140

509 IGPPIGPKBPGLPBPFPFGVKGKGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140

509 IGPPIGPKBPGLPBPFPFGVKGKGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140

509 IGPPIGPKBPGLPBPFPFGVKGKGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140

509 IGPPIGPKBPGLPBPFPFGVKGKGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140

509 IGPPIGPKBPGLPBPFPFGVKGKGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140

509 IGPPIGPKBPGLPBPFPFGVKGKGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140

509 IGPPIGPKBPGLPBPFPFGVKGKGVAGLHGPFGKGLGPPGQGLPBPFPBP 568

141 -----CSCSGHTKSAFSAVATKSYR 162

569 GPPVMPPTPAPDSEYLPDMGLGIDVKTTPAAYAKKKGKGGPAPVEMPAFTAEITAPPP 628

163 ERLPIKEDKILMNEGHNASSGKFCVGPPIYFTYDITLANKHLAGLVHNGQYRI 222

629 VGARIKEDRLLYNRKVNPTGTFTCEVPVYFAVHVHCKGGVWVALFKNNBPVWT 688

223 FDA-NTGNHVASSTILAKQGBEWTQITYSQNGLFYPPYTDLSFTGFLY 276

689 YDEYKKGFLLDQASGSAVALLRPGRVFLQ--PSEDAAGLVAGQYVHSSFSGLY 742

Db

Query

102 IGRAPRGPKNVNGTPCK-----HGTGKKGPKGKKGEPGLP----- 140



## RESULT 13

S29328 Complement subcomponent C1q chain C - mouse

C/Species: Mus musculus (house mouse)

C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C/Accession: S29328

R/Petry, F.; Reid, K.B.M.; Loos, M.

A/Title: Isolation, sequence analysis and characterization of cDNA clones coding for the ecebelin.

A/Reference number: S29328; MUID:9301118; PMID:1396691

A/Accession: S29328

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-246 &lt;PRT&gt;

A/Cross-references: UNIPROT:Q02105; EMBL:X66295; NID:G50228; PIDN:CAA4693.1; PID:G50225  
C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom  
F/122-245/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 24.8%; Score 393; DB 2; Length 246;

Best Local Similarity 38.8%; Pred. No. 1.7e-21;

Matches 94; Conservative 36; Mismatches 84; Indels 28; Gaps 11;

Qy 42 QGPPGPGAPGAPSGMGMGPFPGKDGQDGHDDGDSGEGPPTGNGKPGKPKAGA 101

Db 28 QASAGCGTGTGMPGPM-----PGAPGKDGHGDLGSPKGEPTIPAVPTGPGKQKQKGPBM 81

Qy 102 IGRAGPAPGPGVNGTPEGKGTGPKGKPKGKGGKPGGLPQPCSCSGHTKSAFSA- VTKSY 160

Db 82 PGRHKGKPGKGTGTS-----GLPGDPGPRGPGEGVEGRYK--QKH-QSVFTVTRQTQY 132

Qy 161 PREHLPIKPKILMNEGHTYASGKVCVPGTGYTYDTT-LAN--KHLAIGLVNNGQ 217

Db 133 PEANALRFNSVTVNPGSHVNPSTGKTCVEPGLYFYFYTHANLAVHMLNLA----- 188

Qy 218 YRIKTF-DANTGNHDVAGSTTLAKOGDEWVLQIFYSEONGLFYDPYTDLSLFTGLFY 276

Db 189 -RVASFCDHMFNSKQVSSGGLLRLQKQDEWLVSV--NDVNGM-VGIEGNSVSVSGFLIF 244

Qy 277 AD 278

Db 245 PD 246

## RESULT 14

C1H0QB Complement subcomponent C1q chain B precursor [validated] - human

N/Alternate names: complement subcomponent C1q beta chain

C/Species: Homo sapiens (man)

C/Date: 22-May-1981 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004

C/Accession: B23422; A23422; B90304; A90301; B90315; A03206

R/Reid, K.B.M.

A/Title: Molecular cloning and characterization of the complementary DNA and gene coding

A/Reference number: A23422; MUID:86076906; PMID:3000358

A/Accession: B23422

A/Molecule type: DNA

A/Residues: 'HS', 1-32 &lt;REI&gt;

A/Cross-references: UNIPROT:P02746

A/Note: the authors translated the codon AGT for the second position as Arg; they were u

A/Accession: A23422

A/Molecule type: mRNA

A/Residues: 28-253 &lt;REI&gt;

A/Cross-references: EMBL:X03084

A/Note: the authors translated the codon ACA for residue 46 as Ile

R/Reid, K.B.M.

Biochem. J. 179, 367-371, 1979

A/Title: Complete amino acid sequences of the three collagen-like regions present in sub

A/Reference number: A90304; MUID:80020137; PMID:486087

A/Accession: B90304

A/Molecule type: protein

A/Residues: 'E', 29-84, 'D', 86-99, 'P', 101-135 &lt;RES&gt;

R/Reid, K.B.M.; Thompson, E.O.P.

Biochem. J. 173, 863-868, 1978

A/Title: Amino acid sequence of the N-terminal 108 amino acid residues of the B chain of

A/Reference number: A90301; MUID:79041552; PMID:708376

A/Accession: A90301

A/Molecule type: protein

A/Residues: 28-99, 'P', 101-195 &lt;RES&gt;

R/Reid, K.B.M.; Gagnon, J.; Frampton, J.

Biochem. J. 203, 559-569, 1982

A/Title: Completion of the amino acid sequences of the A and B chains of subcomponent C1

A/Reference number: A90315; MUID:82283890; PMID:6984411

A/Accession: B90315

A/Molecule type: protein

A/Residues: 136-253 &lt;REA&gt;

A/Note: 176-Glx may also be present

C/Comment: The first component of complement is a calcium-dependent complex of the three

activation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.

C/Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfide

(see PIR:C1HUQ) chain. Equimolar amounts of the A, B, and C chains are found after redu

C/Genetics:

A/Genes: GDB:C1QB

A/Cross-references: GDB:119043; OMIM:120570

A/Map position: 1p36.3-1p34.1

C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom

C/Keywords: complement pathway; glycoprotein; heterodimer; hydroxylysine; hydroxyproline

F/1-27/Domain: signal sequence #status predicted &lt;SIG&gt;

F/28-253/Product: complement subcomponent C1q chain B #status experimental &lt;MAT&gt;

F/33-116/Domain: collagenous, triple helix &lt;COL&gt;

F/123-249/Domain: complement C1q carboxyl-terminal homology &lt;C1Q&gt;

F/28/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen

F/31/Disulfide bonds: interchain (to chain A-26) #status experimental

F/35.38.41.53.56.65.83.86.101.104.107/Modified site: 4-hydroxyproline (Pro) #status expe

F/59.62.72.92.98.110/Modified site: 5-hydroxylysine (Lys) #status experimental

F/59.62.98.110/Binding site: carboxylate (Lys) (covalent) #status experimental

Query Match 23.8%; Score 376.5; DB 1; Length 253;

Best Local Similarity 37.2%; Pred. No. 2.8e-20;

Matches 94; Conservative 34; Mismatches 92; Indels 33; Gaps 10;

Qy 33 QLVCS-LPGPGPPGAPGAPSGMGMGPFPGKDGQDGHDDGDSGEGPPTGNGKPGKPKAGA 91

Db 28 QLSCTGPPAIPGIGIPETPGDQPGTPTGKGRGKGLGADHGEKEDPPIGAPG 87

Qy 92 KPGPKAGAIIGRAGPAPGPGVNGTPEGKGTGPKGKPKGKGGKPGGLPQPCSCSGHTKSA 151

Db 88 KVGPK-----GPGPK-----GAPAPGAPGKSGD-----YATQKIA 123

Qy 152 PSVAVTKSY-REKLPIKFDKILMNEGHTYASGKVCVPGTGYTYDTTLANKLAI 210

Db 124 PSATRTINVLPRRDQTRIFDHVITNMNNVEPRSGKFTCKVGLYFYTHAS-SRGNLCV 182

Qy 211 GLVNGQ-YRIKTF-DANTGNHDVAGSTTLAKOGDEWVLQIFYSEONGLFYDPYTD 267

Db 183 NLMGRRAQKVCVPCYAVNTFQVTTGAVLKLKEGENVFLQ--TDKISL- GMEGAN 239

Qy 268 SLPTGLIYADQD 280

Db 240 SIFSGLFPDME 252

## RESULT 15

S19018 Complement subcomponent C1q chain A precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004

C/Accession: S19018

R/Petry, F.; Reid, K.B.M.; Loos, M.

A/Title: Gene expression of the A- and B-chain of mouse C1q in different tissues and the

A/Reference number: S19018; MUID:92043789; PMID:1940381

A/Accession: S19018

A/Molecule type: mRNA

A/Residues: 1-245 &lt;PRT&gt;

A/Cross-references: UNIPROT:P98086; EMBL:X58861; NID:G50226; PIDN:CAA41664.1; PID:G50227

C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom  
F;116-243/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 23.3%; Score 369; DB 2; Length 245;

Best Local Similarity 37.1%; Pred. No. 9.4e-20; Matches 92; Conservative 33; Mismatches 87; Indels 36; Gaps 9;

```
QY 35 VCSLP-GPGPPGPGAPGSPGMMGRMGFPKDG-ODGHDGDRGDSGEGPPGRTGNRGK 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 VCRAPNCKDGAPEGPRGPRGLKGERGEGAGIRTGIRGPKDPGESGPPGKRNVL 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 93 PGPCKAKAIGRAGPRGPKGVNGTYPGKHGTTPKKGPKKKGEPGLPGPCSCSGHTYSAP 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 PGP---SGPLGDSGPGGLKGVKGNPG-----NIRDQ-----RPAF 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 153 SVAVTSYPRERLPKEKDLIMNEGCHYNASSGKFCVCGVPGIYVETVD-ITLANKHLAIG 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 S-AIRQNPMTLGNVVIPIKVLITNQESPYQNHTRFICAVPGFYFNFQVLSKMDLCLFIK 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 LVHNGQYRIRTFDANTGN--HDVASGSTILALKQDEVMQLFYSEONGLFYDPYWTDS 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 SSSGGQPRDLSFSENTNNKGLFQVLAGTVQLARGDENVIE--KDPKAGRIYGTETADS 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 269 LFTGFLIY 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 IFSGFLIF 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: December 25, 2004, 08:39:35  
Job time : 76.0538 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: December 25, 2004, 06:42:17 ; Search time 422.991 Seconds  
(without alignments)  
387.672 Million cell updates/sec

Title: US-10-621-787-2

Perfect score: 1582  
Sequence: 1 MIPWVLLACGLPCADPLG.....TDSLFTGFLIYADDDNEV 285

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Uniprot 02:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1582	100.0	285	1	COT2_HUMAN
2	1492	94.3	294	2	COT2_HUMAN
3	898.5	56.8	312	2	O8CHX9
4	895	56.6	289	1	COT7_HUMAN
5	889	56.2	289	2	O8BVD7
6	521.5	33.0	333	2	O8IUD4
7	521.5	33.0	333	2	AAQ88512
8	516.5	32.6	680	1	CA1A_HUMAN
9	511	32.3	680	1	CA1A_HUMAN
10	503	31.8	295	2	O9ZIK4
11	502.5	31.8	674	1	CA1A_BOVIN
12	502.5	31.8	675	2	O9N178
13	493	31.2	674	1	CA1A_CHICK
14	482	30.5	508	1	COT1_ONCOK
15	457.5	28.9	419	1	COLE_LEPMA
16	444	28.1	703	1	CA28_HUMAN
17	442.5	28.0	743	1	CA18_MOUSE
18	442.5	28.0	744	2	O8BGL6
19	442.5	28.0	744	2	O921S8
20	441.5	27.9	244	2	O6QW57
21	441.5	27.9	244	2	AA567924
22	440	27.8	640	2	O6KAQ4
23	440	27.8	640	2	BAD21403
24	440	27.8	699	2	O6PIC4
25	440	27.8	699	2	AA65148
26	439	27.7	744	1	CA18_RABIT
27	436	27.6	744	1	CA18_HUMAN
28	436	27.6	744	2	AA688919
29	435	27.5	744	2	O71ZK2
30	431.5	27.3	244	1	APM1_HUMAN
31	426	26.9	247	2	O8BRF2

32	425	26.9	240	2	O95MO4	O95mq4 bos taurus
33	425	26.9	247	1	APM1_MOUSE	O60994 mus musculus
34	424	26.8	195	2	O8B2E3	O8B283 mus musculus
35	424	26.8	247	2	AA828770	AA828770 mus musculus
36	420	26.5	243	2	O95JD7	O95jd7 macaca mulia
37	419.5	26.5	235	2	O6DJ45	O6dj45 xenopus tro
38	413	26.1	244	2	O8K3R4	O8K3r4 rattus norv
39	411.5	26.0	235	2	O6GLS9	O6gl59 xenopus lae
40	409	25.9	243	2	O6PP07	O6pp07 sus scrofa
41	409	25.9	243	2	AA100459	AA100459 sus scrofa
42	408	25.8	242	2	O6Q2K6	O6q2k6 sus scrofa
43	408	25.8	242	2	AA575592	AA575592 sus scrofa
44	408	25.8	243	2	O8K479	O8k479 sus scrofa
45	404	25.5	243	2	O8R002	O8r002 mus musculus
46	403	25.5	243	1	COT5_HUMAN	O9bpx10 homo sapien
47	402	25.4	243	2	O8N6P2	O8n6p2 homo sapien
48	400	25.3	246	2	O6DI63	O6di63 mus musculus
49	393	24.8	246	2	C1QC_MOUSE	O02105 mus musculus
50	392	24.8	243	2	O7YRF8	O7yrf8 sus scrofa
51	392	24.8	244	2	O76C76	O76c76 canis famli
52	392	24.8	244	2	BAD15362	BAD15362 canis fam
53	376.5	23.8	251	1	C1QB_HUMAN	P02746 homo sapien
54	369	23.3	245	1	C1QA_MOUSE	P98086 mus musculus
55	369	23.3	245	2	O9DCM6	O9dc6m m mus muscu
56	369	23.3	253	2	AA67001	AA67001 mus muscu
57	367.5	23.2	240	2	O6LA35	O6la35 lampetra ja
58	367	23.2	236	2	O6GNJ3	O6gnj3 xenopus lae
59	366.5	23.2	194	2	O95J95	O95j95 canis famli
60	366	23.1	245	1	C1QC_HUMAN	P02747 homo sapien
61	363.5	23.0	258	1	C1RF_MOUSE	O88992 mus musculus
62	362	22.9	253	1	C1QB_MOUSE	P14106 mus musculus
63	357	22.6	245	2	O7Z502	O7z502 homo sapien
64	357	22.6	253	1	C1QB_RAT	P31721 rattus norv
65	355.5	22.5	258	1	C1RF_HUMAN	O75973 homo sapien
66	355	22.4	247	2	O6GPX5	O6gpx5 xenopus lae
67	348.5	22.0	245	1	C1QA_HUMAN	P02745 homo sapien
68	341	21.6	255	1	C1QC_MOUSE	O9684 mus musculus
69	331	20.9	287	2	O8CFE0	O8cfe0 mus musculus
70	328	20.7	238	2	O86Z23	O86z23 homo sapien
71	327.5	20.7	246	2	O6NZU2	O6nzu2 brachydario
72	327.5	20.7	246	2	AA655968	AA655968 brachydario
73	316	20.0	287	2	O7Z5L3	O7z5l3 homo sapien
74	315.5	19.9	250	2	O7Z282	O7z282 brachydario
75	311.5	19.7	1069	2	O6LA8	O6la8 homo sapien
76	311.5	19.7	1069	2	CA67261	CA67261 homo sapi
77	311.5	19.7	1461	1	CA11_CANPA	O9x6j7 canis famli
78	311.5	19.7	1461	2	O76045	O76045 homo sapien
79	311.5	19.7	1464	1	CA11_HUMAN	P02452 homo sapien
80	311.5	19.7	1464	2	O8N473	O8n473 homo sapien
81	310	19.6	1225	2	O6PCU3	O6pcu3 mus musculus
82	310	19.6	1225	2	AAH5281	AAH5281 mus muscu
83	310	19.6	1453	1	CA11_MOUSE	P11087 mus musculus
84	310	19.6	1453	2	O810J9	O810j9 mus musculus
85	309.5	19.6	187	2	O6L7J6	O6l7j6 cyprinus ca
86	309.5	19.6	187	2	BAD22535	BAD22535 cyprinus
87	309	19.5	671	1	CA11_RAT	P02454 rattus norv
88	308	19.5	1453	2	O63079	O63079 rattus norv
89	306	19.3	215	1	HP25_TAWSI	O65576 tamias sibi
90	306	19.3	1453	1	CA11_CHICK	O9es30 m collageneo
91	305.5	19.3	246	1	COT3_HUMAN	O9bpx14 homo sapien
92	304	19.2	246	2	BAC1676	BAC1676 homo sapi
93	304	19.2	304	2	AAK6452	AAK6452 caenorhabd
94	303.5	19.2	431	2	O95Y22	O95y22 caenorhabd
95	303.5	19.2	779	1	CA11_BOVIN	P02453 bos taurus
96	302.5	19.1	343	2	O46388	O46388 equus cabal
97	302.5	19.0	297	2	O8MXR1	O8mxr1 caenorhabd
98	300.5	19.0	299	2	O22393	O22393 caenorhabd
99	300.5	19.0	717	2	O9NQ52	O9nq52 homo sapien
100	300.5	19.0				

## ALIGNMENTS





RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).

RL [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RA The PANTOM Consortium,

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs." Nature 420:563-573(2002).

RL [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes." Genome Res. 10:1617-1630(2000).

RL [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Komno H., Akiyama J., Nishi K., Kitamura T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujikawa S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "Riken integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer." Genome Res. 10:1757-1771(2000).

RL [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RA Adachi J., Aizawa K., Akahira S., Akiyama T., Arai A., Aono H., Atakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoaka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Komno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numachi R., Ono M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;

RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RL [7]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N-3;

RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., Brownstein M.J., Uedin T.B., Tishiyuki S., Carninci P., Prange C., Rana S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Pheeby J., Helton B., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shvachenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Kozmyński M.I., Skelton U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP [8]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N-3;

RC TISSUE=Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy.

RA Strausberg R.;

RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; AK007683; BAB25187.1;

DR EMBL; BC030324; AAH03024.1;

DR HSSP; Q60994; 1C28.

DR MGD; MG1:1916433; C1qtnf2.

DR GO; GO:0005737; C:cytoplasm; IEA.

DR GO; GO:0006817; P:phosphate transport; IEA.

DR InterPro; IPR001073; C1q.

DR InterPro; IPR008160; C1qagen.

DR InterPro; IPR008983; TWF\_1like.

DR Pfam; PF00386; C1q; 1.

DR Pfam; PF01391; Collagen; 2.

DR PRINTS; PR00007; COMPLEMENTC1Q.

DR SMART; SM00110; C1Q; 1.

DR PROSITE; PS01113; C1Q; 1.

KW Collagen.

SQ SEQUENCE 294 AA; 30865 MW; 6D3905AE7C19E6FA CRC64;

Query Match 94.3%; Score 1492; DB 2; Length 294;

Best Local Similarity 94.4%; Pred. No. 9,1e-92;

Matches 269; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Oy 1 MIPVTLALCALPCAADPLGAFARDPKSGPOLVCSLPFGQGPFGAPGPGMGGRM 60

Db 10 MISVVLALCALPCAADPLGAFARDPKSGPOLVCSLPFGQGPFGAPGPGMGGRM 69

Oy 61 GFPKGDQDGDHGDGSDSGEPPGRTGNRKGPKKAGAIAGAPGPGKVGNGTPEKH 120

Db 70 GFPKGDQDGDGDRGDSGEGPRTGNRKGPKKAGAIAGAPGPGKVGNGTPEKH 129

Oy 121 GTPPKKPKPKKGGKRGRLPGPSCGSGHTKSAFSAVNVKSYREPLPKFDKILNNEGSHY 180

Db 130 GTPPKKPKPKKGGKRGRLPGPSCGSGHTKSAFSAVNVKSYREPLPKFDKILNNEGSHY 189

Oy 181 NASGKRVCGVPGVGYFTYDITLANKHAIAGLVNNGVYRTPDANTGNHVASGSTLLA 240

Db 190 NASGKRVCGVPGVGYFTYDITLANKHAIAGLVNNGVYRTPDANTGNHVASGSTLLA 249

Oy 241 LKQDDEVWLTQIFYSEONGLPFYDPTWDSLFTGFLIYADQDDPNEV 285

Db 250 LKQDDEVWLTQIFYSEONGLPFYDPTWDSLFTGFLIYADQDDPNEV 294

RESULT 3

O8CHX9 PRELIMINARY; PRT; 312 AA.

ID O8CHX9

AC O8CHX9

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE C1qntf7 protein (Fragment).

GN Name=C1qntf7;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI TaxID=10090;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Mammary gland;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,

RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carminci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalka U., Smallus D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
RA Strauberg R.;  
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC038308; AAI38308.1; -  
DR HSSP: O60994; 1C28.  
DR MGD; MGI:1925911; C1qtnf7.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro: IPR001073; C1q.  
DR InterPro: IPR008160; Collagen.  
DR InterPro: IPR008983; TNF-like.  
DR Pfam; PF00386; C1q; 1.  
DR Pfam; PF01391; Collagen; 2.  
DR PRINTS; PR00007; COMPLEMENTC1Q.  
DR SMART; SM00110; C1q; 1.  
DR PROSITE; PS01113; C1q; 1.  
KM Collagen.  
FT NON TMR.  
SQ SEQUENCE 312 AA; 33162 MW; 1E240A83A5014F43 CRC64;  
Query Match 56.8%; Score 898.5; DB 2; Length 312;  
Best Local Similarity 62.7%; Pred. No. 3.4e-52;  
Matches 163; Conservative 32; Mismatches 62; Indels 3; Gaps 1;  
QY 24 RDRPRK---SPQVCSLPPQPPGPPGAPSGMMGRMGFPKQDQDGHDRGDSGR 80  
DB 42 RANQAKESYSPRYICSPGAPGPPGANGSGPGRIGLPDRDRDGKKGKGGKGT 101  
QY 81 EGPPGRTGNGKPKPKKAGAGAGRGPGRGVNGTGPKGKGPCKGKGGKGGPGL 140  
DB 102 AGLKGGTGGPGLAGSGDQGTGKKGPIGPGEGKEGVGPACPPPKDQDQDGLPGV 161  
QY 141 CSCSGHTKSAFSAVATKSYPRERLPFKDKILNBSGHNVAASGKFEVCGVPGIYFTYD 200  
DB 162 CRGSGIVKSAFSAVATKSYPRERLPFKDKILNBSGHNVAASGKFEVCGVPGIYFTYD 221  
QY 201 ITLANKKLAIGLVNCGQRIITPDANTGNHDVAGSTTLAKQDDEWLQIFYSQNGLP 260  
DB 222 ITLANKKLAIGLVNCGQRIITPDANTGNHDVAGSTTLAKQDDEWLQIFYSQNGLP 281  
QY 261 YDPYWTSLFTGFLIYADQD 280  
DB 282 SDPGWADSLFSGFLIYDID 301

OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Piddington C.S., Sheppard P.O., Bishop P., Laesser G.W.;  
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
EX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buétow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carminci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smallus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- SIMILARITY: Contains 1 C1q domain.  
CC -1- SIMILARITY: Contains 1 collagenous domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----  
DR EMBL; AF329839; AAI37963.1; -  
DR EMBL; BC022187; AAI322187.1; -  
DR EMBL; BC024015; AAI24015.1; -  
DR HSSP; O60994; 1C28.  
DR Genew; HGNC:14342; C1QTNF7.  
DR InterPro: IPR001073; C1q.  
DR InterPro: IPR008160; Collagen.  
DR InterPro: IPR008983; TNF-like.  
DR Pfam; PF00386; C1q; 1.  
DR Pfam; PF01391; Collagen; 2.  
DR PRINTS; PR00007; COMPLEMENTC1Q.  
DR SMART; SM00110; C1q; 1.  
DR PROSITE; PS01113; C1q; 1.  
KM Collagen; signal.  
FT SIGNAL 1 16 Potential.  
FT CHAIN 17 289 Complement C1q tumor necrosis factor-  
FT DOMAIN 38 139 related protein 7.  
FT DOMAIN 141 276 Collagen-like.  
SQ SEQUENCE 289 AA; 30683 MW; A61609FF86D26946 CRC64;  
Query Match 56.8%; Score 895; DB 1; Length 289;  
Best Local Similarity 63.6%; Pred. No. 5.4e-52;  
Matches 159; Conservative 33; Mismatches 58; Indels 0; Gaps 0;  
QY 31 SPQVCSLPPQPPGPPGAPSGMMGRMGFPKQDQDGHDRGDSGEGPPGRTGNR 90  
DB 29 SPRTGSGPGLPGRPPGANGSGPGRIGLPDRDRDGKKGKGGKGTAGLRGKTGPL 88  
QY 91 GKPPKKAAGIAGRAGRGKGVNGTGPKGKGTGKPKKKGKGGKGGPGLPGCSCSGHTKS 150  
DB 89 GLAGEKDQGTGKKGPIGPGEGKEGVGPACPPPKDQDQDGLPGVCRGSGIVLKS 148

QY 151 AFSVAVTKSPREBRLPIKEDKILNMGHVNASSGKPGVCGIYYFTYDITLANKHAI 210  
 DB 149 AFSVGITTSYBERPLIFENVLFNEGHYNPATGKPLCAPPGICYFSDITLANKHAI 208  
 QY 211 GLVHNGQYRITFPANTGNHNVASGSTIILAKOGDEWLIQIFYSBONGLFYDPYWTDSL 270  
 DB 209 GLVHNGQYRITFPANTGNHNVASGSTIYIYQPBDEWLEIFPDQNLFPDPCGADSLF 268  
 QY 271 TGFLLIYADQ 280  
 DB 269 SGFLLYVDT 278

RESULT 5  
 Q8BVD7 PRELIMINARY; PRT; 289 AA.  
 ID Q8BVD7  
 AC Q8BVD7;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Mus musculus 16 days embryo lung CDNA, RIKEN full-length enriched library, clone:8430425G24 product:COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-RELATED PROTEIN homolog.  
 DE Name=C1qtnf7;  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Mech. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA The FANTOM Consortium;  
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team; Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Komano H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Azawa K., Nagaoka S., Sasaki N., Carninci P., Komano H., Akiyama J., Nishi K., Kikunishi T., Tashiro H., Itoh M., Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiwaagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA Adachi J., Azawa K., Akahira S., Akimura T., Aono H., Arai A., Arahawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Komano H., Kouda M., Koye S., Kurihara C., Matsuura T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H., Shibata K., Shibata Y., Shinagawa A., Tanaka T., Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;  
 RA Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.  
 RL EMBL; AK078818; BAC37409.1; -.  
 DR HSSP; Q60994; IC28.  
 DR MGD; MGI:1925911; C1qtnf7.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR001073; C1q.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008983; TNF-like.  
 DR Pfam; PF00386; C1q; 1.  
 DR Pfam; PF01391; Collagen; 2.  
 DR PRINTS; PR00007; COMPLEMENTC1Q.  
 DR SMART; SM00110; C1Q; 1.  
 DR KEGG; C00000; C1Q.  
 SQ SEQUENCE 289 AA; 30483 MW; 514AC7FC318832A CRC64;

Query Match 56.2%; Score 889; DB 2; Length 289;  
 Best local similarity 63.6%; Pred. No. 1.3e-51;  
 Matches 159; Conservative 31; Mismatches 60; Indels 0; Gaps 0;

QY 31 SPOLVCSLPQPGPGPGAPGAPGSGMMGMGFPKDGQDGHGDSGSEGGPPGRTGNR 90  
 DB 29 SPRYICSTPLGPGPPGPGANGSPGPHGRIGLPDRDGDGKSGKSGKGTGLGKGTPL 88  
 QY 91 GKPPKKGAGIAGAPRGPGVNGTPEKIGTPEKKGKPKKKGKGBGLFGPCSCSGATKS 150  
 DB 89 GLAGEKGDQSTGKGPFGPEGKEVGVPAPPPGKGRDQDGLFGVCRCSIVLKS 148  
 QY 151 AFSVAVTKSPREBRLPIKEDKILNMGHVNASSGKPGVCGIYYFTYDITLANKHAI 210  
 DB 149 AFSVGITTSYBERPLIFENVLFNEGHYNPATGKPLCAPPGICYFSDITLANKHAI 208  
 QY 211 GLVHNGQYRITFPANTGNHNVASGSTIILAKOGDEWLIQIFYSBONGLFYDPYWTDSL 270  
 DB 209 GLVHNGQYRITFPANTGNHNVASGSTIYIYQPBDEWLEIFPDQNLFPDPCGADSLF 268  
 QY 271 TGFLLIYADQ 280  
 DB 269 SGFLLYVDT 278

RESULT 6  
 Q8IUU4 PRELIMINARY; PRT; 333 AA.  
 ID Q8IUU4  
 AC Q8IUU4;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Hypothetical protein WGC48915 (R1W6503).  
 GN Name=WGC48915; ORFName=UNQ06503;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932;

RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatzenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinck P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzyzanski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences." J.  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RN  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RL Straubeberg R.;  
 RL Submitted (NOV-2002) to the EMBL/Genbank/DBJ databases.  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22887296; PubMed=12975309;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Helens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seehagiri S., Simmons L., Singh V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RT "The secreted protein discovery initiative (SPDI), a large-scale  
 RT effort to identify novel human secreted and transmembrane proteins: a  
 RT bioinformatics assessment." J.  
 RL Genome Res. 13:2265-2270(2003).  
 RX EMBL; BC040438; AA040438.1; -;  
 DR EMBL; AY358145; AA088512.1; -;  
 DR HSSP; Q60994; 1C28.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphatase transport; IEA.  
 DR InterPro; IPR001073; Clq.  
 DR InterPro; IPR008161; Clq\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008983; TWF\_like.  
 DR Pfam; PF00386; Clq; 1.  
 DR Pfam; PF01391; Collagen; 3.  
 DR PRINTS; PR00007; COMPLEMENT1Q.  
 DR PRODOM; PD000007; Clq\_helix; 1.  
 DR PROSITE; PS01113; Clq; 1.  
 KW Collagen; Hypothetical protein.  
 SQ SEQUENCE 333 AA; 34649 MW; E9EBF803A034E307 CRC64;

Query Match 33.0%; Score 521.5; DB 2; Length 333;  
 Best Local Similarity 38.3%; Pred. No. 4,7e-27;  
 Matches 128; Conservative 37; Mismatches 102; Indels 67; Gaps 11;

QY 4 WLLACALPCADPLGAFARRD-FRKSPOLVCS-----LPGQGPFGPAPGAPSGMM 57  
 DB 5 WLLATIRI-CT-----GNINSQDTCRQGHGPGIPGNPGHNLPGRDGDAKGDAGDEP 58

QY 58 GRMGFPKGDQDGHGDRGDS-----GEE 81  
 DB 59 GRPSPKDGSGKGRGADGKVEAKGIGKDGSRGSPGKGPGLAGPWEKGLGET 118

QY 82 GPGKRTGNRGPCKGKAGAIGRAPRGKGVNGTPGKHGTPGK--GPKKGGEGGLP 138  
 DB 119 GPGQKGNKNDVGPTEGEGPRGNIGPLGPTGLPQPMWPIGKPGKGEAGPTGPGSEBGR 178

QY 139 GPGSCGS-----GHT-----KSAFSAVT--KSYPERLRPIKFDKILNMGHYNASS 184  
 DB 179 GIRKWKDDBRKGKIGTTLVLPKSAFTVGLTWSKSPDPIKFDKILNMFHNYDTAA 238

QY 185 GKFCVGVPIGYFTYDITLANKHLAIGLVNNGQYRIRTPDANTGNHVASGSTITALKOG 244  
 DB 239 GKFTCHLAGYVYFTYHTVPSRNVQSLVNGGVILHTKQAYMSSEDAAGIVLQKLG 298

QY 245 DEWVLQIFYSEQ-NGLFYDPYWDLSFTGFLIYA 277  
 DB 299 DEWVLQVTGGERFNGLPADBD-DDTPTGFLIFS 331

RESULT 7  
 ID AA088512 PRELIMINARY; PRT; 333 AA.  
 AC AA088512;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE R1W6503.  
 GN UNG6503.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primata; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12975309;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Helens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seehagiri S., Simmons L., Singh V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RT "The Secreted Protein Discovery Initiative (SPDI), a Large-Scale  
 RT Effort to Identify Novel Human Secreted and Transmembrane Proteins: A  
 RT Bioinformatics Assessment." J.  
 RL Genome Res. 13:2265-2270(2003).  
 DR EMBL; AY358145; AA088512.1; -;  
 SQ SEQUENCE 333 AA; 34649 MW; E9EBF803A034E307 CRC64;

Query Match 33.0%; Score 521.5; DB 2; Length 333;  
 Best Local Similarity 38.3%; Pred. No. 4,7e-27;  
 Matches 128; Conservative 37; Mismatches 102; Indels 67; Gaps 11;

QY 4 WLLACALPCADPLGAFARRD-FRKSPOLVCS-----LPGQGPFGPAPGAPSGMM 57  
 DB 5 WLLATIRI-CT-----GNINSQDTCRQGHGPGIPGNPGHNLPGRDGDAKGDAGDEP 58

QY 58 GRMGFPKGDQDGHGDRGDS-----GEE 81  
 DB 59 GRPSPKDGSGKGRGADGKVEAKGIGKDGSRGSPGKGPGLAGPWEKGLGET 118

QY 82 GPGKRTGNRGPCKGKAGAIGRAPRGKGVNGTPGKHGTPGK--GPKKGGEGGLP 138  
 DB 119 GPGQKGNKNDVGPTEGEGPRGNIGPLGPTGLPQPMWPIGKPGKGEAGPTGPGSEBGR 178

QY 139 GPGSCGS-----GHT-----KSAFSAVT--KSYPERLRPIKFDKILNMGHYNASS 184  
 DB 179 GIRKWKDDBRKGKIGTTLVLPKSAFTVGLTWSKSPDPIKFDKILNMFHNYDTAA 238

QY 185 GKFCVGVPIGYFTYDITLANKHLAIGLVNNGQYRIRTPDANTGNHVASGSTITALKOG 244  
 DB 239 GKFTCHLAGYVYFTYHTVPSRNVQSLVNGGVILHTKQAYMSSEDAAGIVLQKLG 298

QY 245 DEWVLQIFYSEQ-NGLFYDPYWDLSFTGFLIYA 277  
 DB 299 DEWVLQVTGGERFNGLPADBD-DDTPTGFLIFS 331

RESULT 8  
 CALI\_HUMAN STANDARD; PRT; 680 AA.  
 ID CAIA\_HUMAN  
 AC 003692;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Collagen alpha 1(X) chain precursor.  
 GN Name=COL10A1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92109659; PubMed=1764025;  
 RA Thomas J.T., Cresswell C.J., Raab B., Nicolai H., Jones T.,  
 RA Solomon E., Grant M.E., Boot-Handford R.P.;  
 RT "The human collagen X gene. Complete primary translated sequence and  
 RT chromosomal localization.";  
 RL Biochem. J. 280:617-623(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93012005; PubMed=1397333;  
 RA Reichenberger E., Beter F., Luvallie P., Olsen B.R., von der Mark K.,  
 RA Bertling W.M.;  
 RT "genomic organization and full-length cDNA sequence of human collagen  
 RT X.";  
 RL FEBS Lett. 311:305-310(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX Beter F., Lammli M.B., von der Mark K.;  
 RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX Williams S.;  
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 52-680 FROM N.A.  
 RX MEDLINE=92267014; PubMed=1587271;  
 RA Apte S.S., Seldin M.F., Hayaishi M., Olsen B.R.;  
 RT "Cloning of the human and mouse type X collagen genes and mapping of  
 RT the mouse type X collagen gene to chromosome 10.";  
 RL Eur. J. Biochem. 206:217-224(1992).  
 RN [6]  
 RP SEQUENCE OF 561-666 FROM N.A.  
 RX MEDLINE=91243838; PubMed=2037056;  
 RA Apte S., Mattei M.-G., Olsen B.R.;  
 RT "Cloning of human alpha 1(X) collagen DNA and localization of the  
 RT COL10A1 gene to the q21-q22 region of human chromosome 6.";  
 RL FEBS Lett. 282:393-396(1991).  
 RN [7]  
 RP SEQUENCE OF 547-655 FROM N.A.  
 RX MEDLINE=92077285; PubMed=1743401;  
 RA Reichenberger E., Aigner T., von der Mark K., Stoeb H., Bertling W.;  
 RT "In situ hybridization studies on the expression of type X collagen in  
 RT fetal human cartilage.";  
 RL Dev. Biol. 148:562-572(1991).  
 RN [8]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=97255959; PubMed=9101290;  
 RA Kivianen H., Tromp G., Prockop D.J.;  
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-  
 RT associated collagen (type IX), and network-forming collagen (type X)  
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";  
 RL Hum. Mutat. 9:300-315(1997).  
 RN [9]  
 RP VARIANTS SMCD ASP-598 AND PRO-614.  
 RX MEDLINE=94134476; PubMed=8304336;  
 RA Wallis G.A., Raab B., Sweetman W.A., Thomas J.T., Super M., Evans G.,  
 RA Grant M.E., Boot-Handford R.P.;

RT "Amino acid substitutions of conserved residues in the carboxyl-  
 RT terminal domain of the alpha 1(X) chain of type X collagen occur in  
 RT two unrelated families with metaphyseal chondrodysplasia type  
 RT Schmid.";  
 RL Am. J. Hum. Genet. 54:169-178(1994).  
 RN [10]  
 RP VARIANTS SMCD ARG-591.  
 RX MEDLINE=94227470; PubMed=8004099;  
 RA McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francomano C.A.;  
 RT "Additional mutations of type X collagen confirm COL10A1 as the Schmid  
 RT metaphyseal chondrodysplasia locus.";  
 RL Hum. Mol. Genet. 3:303-307(1994).  
 RN [11]  
 RP VARIANTS SMCD VAL-618.  
 RX MEDLINE=95181449; PubMed=7876225;  
 RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;  
 RT "Type X collagen multimer assembly in vitro is prevented by a Gly618  
 RT to Val mutation in the alpha 1(X) NCI domain resulting in Schmid  
 RT metaphyseal chondrodysplasia.";  
 RL J. Biol. Chem. 270:4558-4562(1995).  
 RN [12]  
 RP VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.  
 RX MEDLINE=95331767; PubMed=7607655;  
 RA Bonaventure J., Chamande F., Maroteaux P.;  
 RT "Mutations in three subdomains of the carboxy-terminal region of  
 RT collagen type X account for most of the Schmid metaphyseal  
 RT dysplasias.";  
 RL Hum. Genet. 96:58-64(1995).  
 RN [13]  
 RP VARIANTS SMCD PRO-600.  
 RX MEDLINE=96375754; PubMed=8782043;  
 RA Wallis G.A., Raab B., Sykes B., Bonaventure J., Maroteaux P.,  
 RA Zabel B., Wyne-Davies R., Grant M.E., Boot-Handford R.P.;  
 RT "Mutations within the gene encoding the alpha 1(X) chain of type X  
 RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but  
 RT not several other forms of metaphyseal chondrodysplasia.";  
 RL J. Med. Genet. 33:450-457(1996).  
 RN [14]  
 RP VARIANTS SMCD GLU-18 AND ARG-18.  
 RX MEDLINE=97220591; PubMed=9067753;  
 RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;  
 RT "Mutations in the N-terminal globular domain of the type X collagen  
 RT gene (COL10A1) in patients with Schmid metaphyseal chondrodysplasia.";  
 RL Hum. Mutat. 9:131-135(1997).  
 RN [15]  
 RP VARIANTS SMD GLU-595.  
 RX MEDLINE=99057503; PubMed=9837818;  
 RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,  
 RA Nakamura Y.;  
 RT "Mutation of the type X collagen gene 'COL10A1' causes  
 RT spondylometaphyseal dysplasia.";  
 RL Am. J. Hum. Genet. 63:1659-1662(1998).  
 RN [16]  
 RP VARIANTS SMCD CYS-597.  
 RX MEDLINE=99069781; PubMed=9852679;  
 RA Sawai H., Ida A., Nakata Y., Koyama K.;  
 RT "Novel missense mutation resulting in the substitution of tyrosine by  
 RT cysteine at codon 597 of the type X collagen gene associated with  
 RT Schmid metaphyseal chondrodysplasia.";  
 RL J. Hum. Genet. 43:259-261(1998).  
 CC -1- FUNCTION: Type X collagen is a product of hypertrophic  
 CC chondrocytes and has been localized to presumptive  
 CC mineralization zones of hyaline cartilage.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- PTM: Prolines at the third position of the tripeptide repeating  
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
 CC -1- DISBASB: Defects in COL10A1 are the cause of Schmid type  
 CC metaphyseal chondrodysplasia (SMCD) [MIM:156500]. SMCD is a  
 CC dominantly inherited disorder of the osseous skeleton. The  
 CC cardinal features of the phenotype are mild short stature, coxa  
 CC vara and a waddling gait. Radiography usually shows sclerosis of  
 CC the ribs, flaring of the metaphyses, and a wide irregular growth  
 CC plate, especially of the knees.

CC -1- DISEASE: Defects in COL10A1 are a cause of spondylometaphyseal  
 CC dysplasia Japanese type (SMD) [MIM:120110]. SMD comprises a  
 CC heterogeneous group of heritable skeletal dysplasias characterized  
 CC by modifications of the vertebral bodies of the spine and  
 CC metaphyses of the tubular bones.  
 CC -1- SIMILARITY: Contains 1 C1q domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X60382; CAA42933.1; -;  
 DR EMBL; X65120; CAA46236.1; -;  
 DR EMBL; X98568; CAA67178.1; -;  
 DR EMBL; AL121963; CAB87590.1; -;  
 DR EMBL; S68531; AAC60615.1; -;  
 DR EMBL; M74050; AAA61221.1; -;  
 DR EMBL; X72579; CAA51170.1; -;  
 DR EMBL; X72580; CAA51170.1; JOINED.  
 DR PIR; S26396; CGH01D.  
 DR PDB; 1GR3; X-ray; A=521-680.  
 DR Genew; HGNC:2185; COL10A1.  
 DR MIM; 120110; -;  
 DR MIM; 156500; -;  
 DR MIM; 184250; -;  
 DR GO; GO:0005581; C:collagen; TAS.  
 DR GO; GO:0001501; P:skeletal development; TAS.  
 DR InterPro; IPR001073; C1q.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR009883; TNP-like.  
 DR Pfam; PF00386; C1q; 1.  
 DR Pfam; PF01391; Collagen; 8.  
 DR PRINTS; PR00007; COMPLEMENTC1Q.  
 DR PROSITE; PS01113; C1Q; 1.  
 KW 3D-structure; Collagen; Connective tissue; Disease mutation;  
 KW Extracellular matrix; Hydroxylation; Polymorphism; Repeat; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 680 Collagen alpha 1(X) chain.  
 FT DOMAIN 19 56 Nonhelical region (NC2).  
 FT DOMAIN 57 519 Triple-helical region.  
 FT DOMAIN 520 680 Nonhelical region (NC1).  
 FT DOMAIN 545 680 C1q.  
 FT VARIANT 18 18 G-> E (in SMD).  
 FT VARIANT 18 18 /FTID=VAR 001838.  
 FT VARIANT 18 18 G-> R (in SMD).  
 FT VARIANT 18 18 /FTID=VAR 001839.  
 FT VARIANT 545 545 G-> R (in dbSNP:2228547).  
 FT VARIANT 545 545 /FTID=VAR 001840.  
 FT VARIANT 591 591 C-> R (in SMD).  
 FT VARIANT 591 591 /FTID=VAR 001841.  
 FT VARIANT 595 595 G-> E (in SMD and SMD).  
 FT VARIANT 595 595 /FTID=VAR 001842.  
 FT VARIANT 597 597 Y-> C (in SMD).  
 Query Match 32.6%; Score 516.5; DB 1; Length 680;  
 Best Local Similarity 41.5%; Pred. No. 2,1e-26;  
 Matches 120; Conservative 29; Mismatches 93; Indels 47; Gaps 8;

Db 510 PGPPGPGQAVMEEGFIKAGQREPLSGTPLVSANOGVTGMPVSAFTVILSKAPALGTPI 569  
 Qy 168 KPDITLMEGCHYNNASGKGVCPGPIYFTYITLANKHIALGLVHNGQYRIRTPDANT 227  
 Db 570 PFDITLNRQOHDPRTGTYTCQIPGLYSHVHAKGHWVGLYKNGPWPMTTDEYT 629  
 Qy 228 -GNHDVASGSTITALKQGDENVWLQIFYSQNGLFYDPYWTDSLFTGELI 275  
 Db 630 KGVLDQASSGAILIDLTENDQVWLQLPVAESNGLYSSRY-VHSSFSGLV 677  
 RESULT 9  
 CALA MOUSE  
 ID CALA MOUSE STANDARD; PRT; 680 AA.  
 AC 005306;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Collagen alpha 1(X) chain precursor.  
 GN Name=Col10a1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=93143676; PubMed=8424763;  
 RA Elima K., Berola I., Kovati R., Metsaranta M., Garofalo S., Perala M.,  
 RA de Crombrughe B., Vuorio E.;  
 RT "The mouse collagen X gene: complete nucleotide sequence, exon  
 RT structure and expression pattern.";  
 RL Biochem. J. 289:247-253 (1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv; TISSUE=Liver;  
 RX MEDLINE=93238750; PubMed=8477738;  
 RA Kong R.Y.C., Kwan K.M., Lau E.T., Thomas J.T., Boot-Handford R.P.,  
 RA Grant M.B., Cheah K.S.E.;  
 RT "Intron-exon structure, alternative use of promoter and expression of  
 RT the mouse collagen X gene, Col10a-1.";  
 RL Eur. J. Biochem. 213:99-111 (1993).  
 RN [3]  
 RP SEQUENCE OF 51-680 FROM N.A.  
 RC STRAIN=DBA/2J;  
 RX MEDLINE=92267014; PubMed=1587271;  
 RA Apce S.S., Seidin M.F., Hayashi M., Olsen B.R.;  
 RT "Cloning of the human and mouse type X collagen genes and mapping of  
 RT the mouse type X collagen gene to chromosome 10.";  
 RL Eur. J. Biochem. 206:217-224 (1992).  
 RN [4]  
 RP SEQUENCE OF 385-627 FROM N.A.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=92182017; PubMed=1543751;  
 RA Elima K., Metsaranta M., Kallio J., Perala M., Berola I.,  
 RA Garofalo S., de Crombrughe B., Vuorio E.;  
 RT "Specific hybridization probes for mouse alpha 2(I)X and alpha 1(X)  
 RT collagen mRNAs.";  
 RL Biochim. Biophys. Acta 1130:78-80 (1992).  
 CC -1- FUNCTION: Type X collagen is a product of hyperthrophic  
 CC chondrocytes and has been localized to presumptive  
 CC mineralization zones of hyaline cartilage.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- PM: Prolines at the third position of the tripeptide repeating  
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
 CC -1- SIMILARITY: Contains 1 C1q domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>



```

RESULT 11
CA1A_BOVIN
ID CA1A_BOVIN STANDARD; PRT; 674 AA.
AC P23206;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN Name=COL10A1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cartilage;

```

RX MEDLINE=9111313; PubMed=1703407;  
 RA Thomas J.T., Kwan A.P.L., Grant M.B., Boot-Handford R.P.;  
 RT Isolation of cDNAs encoding the complete sequence of bovine type X  
 collagen. Evidence for the condensed nature of mammalian type X  
 collagen genes.  
 RL Biochem. J. 273:141-148(1991).  
 CC -1- FUNCTION: Type X collagen is a product of hypertrophic  
 chondrocytes and has been localized to presumptive  
 mineralization zones of hyaline cartilage.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- PM: Prolines at the third position of the tripeptide repeating  
 unit (G-X-Y) are hydroxylated in some or all of the chains.  
 CC -1- SIMILARITY: Contains 1 C1q domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X53556; CA37624.1; -.  
 DR PIR; S13301; S13301.  
 DR HSSP; Q03692; IGR3.  
 DR InterPro; IPR001073; C1q.  
 DR InterPro; IPR008161; C1q helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008983; TNF-like.  
 DR Pfam; PF00386; C1q; 1.  
 DR Pfam; PF01391; Collagen; 9.  
 DR PRINTS; PR00007; COMPLEMENTC1Q.  
 DR PRODOM; PD000007; C1q helix; 2.  
 DR PROSITE; PS01113; C1Q; 1.  
 KW Collagen; Connective tissue; Extracellular matrix; Glycoprotein;  
 KW Hydroxylation; Repeat; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 674 Collagen alpha 1(X) chain.  
 FT DOMAIN 19 56 Nonhelical region (NC2).  
 FT DOMAIN 57 519 Triple-helical region.  
 FT DOMAIN 520 674 Nonhelical region (NC1).  
 FT DOMAIN 539 674 C1q.  
 FT DISULFID 194 197 By similarity.  
 FT MOD\_RES 460 460 Hydroxyproline (By similarity).  
 FT MOD\_RES 463 463 Hydroxyproline (By similarity).  
 FT CARBOHYD 611 611 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 674 AA; 65546 MW; CD4CA73A03E004CA CRC64;

Query Match 31.8%; Score 502.5; DB 1; Length 674;  
 Best Local Similarity 41.8%; Pred. No. 1.7e-25;  
 Matches 117; Conservative 26; Mismatches 94; Indels 41; Gaps 8;

QY 29 KGSFOLVCSLPFGPP---GPPAPGSGMGMGPPGKDGODGDHGDSDGSEBPPG 85  
 DB 400 KGNP---GLPGRKGDGPIAGSPGLPVPVGPAGAKGVPGHNGEAGPRGVPGLPSTRPPIG 455  
 QY 86 RTGNRGKPRGPKGAKGAIGRAPRGPKG--VNGTPGKHGTGPKGPKKKEBPGI-RGP--- 140  
 DB 456 PPGIIPGPGSKD---VGTPEPPAGIAYKGLNGLPGRPPPPKRNAGBPGI-RGP 512  
 QY 141 -----CSCGSGHT---KSAPSVAVTKSPRRRLPIKFDKILMNE 176  
 DB 513 PGPFGVALPEDFVKAGRPVSNQGVTMGPVSAFTVILSKAPAIAGTPIPFKILYNK 572  
 QY 177 GGHYNASSGKFGVGVPGIYFTYDITLANKHLAIGLVHNGQYRIPTDAN-TGNHDVAGS 235  
 DB 573 QOHYDPTGTGFTCKIPGIIYFYSYHINVKGTAWGLYKNGPWWYTYDEVYIKGYLDASG 632  
 QY 236 STIALKQGDVWLQIYSEONGLFYDPYWTDSLFTGFLI 275  
 DB 633 SAVIDLTENDQVWLQLPNAGSNGI-YSPRYVHSSFSGLV 671

## RESULT 12

ID 09N178 PRELIMINARY; PRT; 675 AA.  
 AC 09N178;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Type X collagen.  
 GN Name=COL10A1;  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
 OC NCBI\_TaxID=9823;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=21015405; PubMed=11130976;  
 RA Nielsen V.H., Bendixen C., Ambjerg J., Sorensen C.M., Jensen H.E.,  
 RA Shukri N.M., Thomsen B.;  
 RT "Abnormal growth plate function in pigs carrying a dominant mutation  
 in type X collagen."  
 RT Mamm. Genome 11:1087-1092(2000).  
 RL EMBL; AF222861; AAF37271.1; -.  
 DR HSSP; AF222861; AAF37271.1; -.  
 DR HSSP; Q03692; IGR3.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR001073; C1q.  
 DR InterPro; IPR008161; C1q helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008983; TNF-like.  
 DR Pfam; PF00386; C1q; 1.  
 DR Pfam; PF01391; Collagen; 9.  
 DR PRINTS; PR00007; COMPLEMENTC1Q.  
 DR PRODOM; PD000007; C1q helix; 3.  
 DR SMART; SM00110; C1Q; 1.  
 DR PROSITE; PS01113; C1Q; 1.  
 KW Collagen.

SQ SEQUENCE 675 AA; 65447 MW; 26397B10310383F9 CRC64;  
 Query Match 31.8%; Score 502.5; DB 2; Length 675;  
 Best Local Similarity 42.5%; Pred. No. 1.7e-25;  
 Matches 119; Conservative 23; Mismatches 97; Indels 41; Gaps 8;

QY 29 KGSFOLVCSLPFGPP---GPPAPGSGMGMGPPGKDGODGDHGDSDGSEBPPG 85  
 DB 401 KGNP---GLPGRKGDGPIAGSPGLPVPVGPAGAKGVPGHNGEAGPRGVPGLPSTRPPIG 456  
 QY 86 RTGNRGKPRGPKGAKGAIGRAPRGPKG--VNGTPGKHGTGPKGPKKKEBPGI-RGP--- 140  
 DB 457 PPGIIPGPGSKD---GNPGRPPAGIAYKGLNGLPGRPPPPKRNAGBPGI-RGP 513  
 QY 141 -----CSCGSGHT---KSAPSVAVTKSPRRRLPIKFDKILMNE 176  
 DB 514 PGPFGVALPEGVKAGRPVSNQGVTMGPVSAFTVILSKAPAIAGTPIPFKILYNK 573  
 QY 177 GGHYNASSGKFGVGVPGIYFTYDITLANKHLAIGLVHNGQYRIPTDAN-TGNHDVAGS 235  
 DB 574 QOHYDPTGTGFTCKIPGIIYFYSYHINVKGTAWGLYKNGPWWYTYDEVYIKGYLDASG 633  
 QY 236 STIALKQGDVWLQIYSEONGLFYDPYWTDSLFTGFLI 275  
 DB 634 SAVIDLTENDQVWLQLPNAGSNGI-YHSSFSGLV 672

## RESULT 13

ID CA1A CHICK STANDARD; PRT; 674 AA.  
 AC P08125;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Collagen alpha 1(X) chain precursor.  
 GN Name=COL10A1;  
 OS Gallus gallus (Chicken).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 CC Gallus  
 CC NCBI\_TaxID=9031;  
 RN [1]  
 RN SEQUENCE OF 48-674 FROM N.A., AND SEQUENCE OF 103-117 AND 453-466.  
 RX MEDLINE=86168227; PubMed=3082876;  
 RA Nimomyia Y., Gordon M., van der Rest M., Schmid T., Linsenmayer T.,  
 RA Olsen B.R.;  
 RT "The developmentally regulated type X collagen gene contains a long  
 RT open reading frame without introns."  
 RL J. Biol. Chem. 261:5041-5050(1986).  
 RN [2]  
 RN SEQUENCE OF 1-75 FROM N.A.  
 RX MEDLINE=89054019; PubMed=2461368;  
 RA Luvall P., Nimomyia Y., Rosenblum N.D., Olsen B.R.;  
 RT "The type X collagen gene. Intron sequences split the 5'-untranslated  
 RT region and separate the coding regions for the non-collagenous amino-  
 RT terminal and triple-helical domains";  
 RL J. Biol. Chem. 263:18378-18385(1988).  
 RN [3]  
 RN REVISIONS TO C-TERMINUS.  
 RX MEDLINE=89380199; PubMed=2476437;  
 RA Yamauchi N., Benya P.D., van der Rest M., Nimomyia Y.;  
 RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs  
 RT demonstrate that type VIII collagen is a short chain collagen and  
 RT contains triple-helical and carboxyl-terminal non-triple-helical  
 RT domains similar to those of type X collagen."  
 RL J. Biol. Chem. 264:16022-16029(1989).  
 CC -1- FUNCTION: Type X collagen is a product of hypertrophic  
 CC chondrocytes and has been localized to presumptive  
 CC mineralization zones of hyaline cartilage.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- PTM: Prolines at the third position of the tripeptide repeating  
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
 CC -1- SIMILARITY: Contains 1 C1q domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; M13496; AAA48736.1; ALT\_SEQ.  
 DR EMBL; J04194; AAA48634.1; -.  
 DR PIR; S23297; S23297.  
 DR HSP; Q03692; ICR3.  
 DR InterPro; IPR001073; C1q.  
 DR InterPro; IPR008161; C1q\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008983; TWF\_like.  
 DR Pfam; PF00386; C1q; 1.  
 DR Pfam; PF01391; Collagen; 8.  
 DR PRINTS; PR00007; COMPLEMENT1Q.  
 DR PRODOM; PD000007; C1q\_helix; 1.  
 DR PROSITE; PS01113; C1Q; 1.  
 KM Collagen; Connective tissue; Direct protein sequencing;  
 KM Extracellular matrix; Hydroxylation; Repeat; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 674 Collagen alpha 1(X) chain.  
 FT DOMAIN 19 52 Nonhelical region (NC2).  
 FT DOMAIN 53 512 Triple-helical region.  
 FT DOMAIN 539 674 Nonhelical region (NC1).  
 FT MOD\_RSS 453 453 C1q.  
 FT MOD\_RSS 456 456 Hydroxyproline.  
 SQ SEQUENCE 674 AA; 66434 MW; EAB48B1EF174B145 CRC64;  
 Query Match 31.2%; Score 493; DB 1; Length 674;  
 Best Local Similarity 35.0%; Pred. No. 7,5e-25;  
 Matches 117; Conservative 33; Mismatches 92; Indels 92; Gaps 9;

QY 32 POLVCSLPGPQPPPGAPGPP-----SGMGMGPPGKQGDGDGHDGDRDSSGSEEPBG 85  
 DB 338 PQLKGLPGENGLGPPKDMGPPVGPAPGPAKGRGRLPGLDKRGYPRGQLPBPKGHPG 397  
 QY 86 RTGNRGK-----PGP-----KKGAGAGAGPPGKGVNGT-----PKHGTPG 124  
 DB 398 LPQKQKDTGAGHGLGPPVGPQGVKGVPGVNGSRGRRGSGIGVAGVPIGPPMPAPG 457  
 QY 125 KK-----GPKKKKGPPGPP-----CSCGSG 146  
 DB 458 AKGAGAPGLPAPAGIVTKGLRPMGLGPPGPKGNGGEPGLPBPBPBPBGOSTIPEG 517  
 QY 147 HTK-----SAPSVATSYRERLPIEDKLINMGGHYN 181  
 DB 518 YVKESEBELSGSMFKAGANQALGMPVSAFTVLISKAYPATVPIKDKILYRQGHYD 577  
 QY 182 ASGKFGVGVGYFFYDITLANKHLAIGLVHNGQVRIPTDA-NTGNHVDVAGSTILA 240  
 DB 578 PRGIFTCRIPGLYFSTVHAKGTNWVALYKNGSPVMTYDYKQYLDQASGAVID 637  
 QY 241 LKQGEVWLQIFYSQNGLFYDPYMTSLFTGFL 274  
 DB 638 LMENDQWVLQLPNSNSGLYSSEX-VHSSFGFL 670  
 RESULT 14  
 OTOL ONCKE STANDARD; PRT; 508 AA.  
 AC P83371;  
 ID 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Otolin-1 precursor.  
 GN Name=OTOL;  
 OS Oncorhynchus keta (Chum salmon).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 CC Proclanchopeterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 CC NCBI\_TaxID=8018;  
 RN [1]  
 RN SEQUENCE FROM N.A., SEQUENCE OF 26-40 AND 452-482, FUNCTION, AND  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=21845944; PubMed=11856329;  
 RA Murayama E., Takagi Y., Ohira T., Davis J.G., Greene M.I.,  
 RA Nagasawa H.;  
 RT "Fish otolith contains a unique structural protein, otolin-1";  
 RL Eur. J. Biochem. 269:688-696(2002).  
 CC -1- FUNCTION: May be part of the internal framework of the otolith  
 CC where it may provide nucleation sites to facilitate calcification.  
 CC -1- TISSUE SPECIFICITY: Selectively expressed in the sacculus where it  
 CC is localised to the otolith, the gelatinous layer of the otolithic  
 CC membrane, and part of the transitional epithelium.  
 CC -1- PTM: N-glycosylated.  
 CC -1- SIMILARITY: Contains 1 C1q domain.  
 CC -1- SIMILARITY: Contains 1 collagenous domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; AB067770; BAB84561.1; -.  
 DR HSP; Q60994; IC28.  
 DR GO; GO:0005576; C:extracellular; IDA.  
 DR GO; GO:0045299; P:otolith mineralization; NAS.  
 DR InterPro; IPR001073; C1q.  
 DR InterPro; IPR008161; C1q\_helix.  
 DR InterPro; IPR008160; Collagen.

```

DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 4.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 1.
DR SMART; SM00110; C1q; 1.
DR PROSITE; PS01113; C1Q; 1.
DR Collagen; Direct protein sequencing; Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 508
FT DOMAIN 144 367
FT DOMAIN 369 505
FT CARBOHYD 121 121
FT CARBOHYD 416 416
SQ SEQUENCE 508 AA; 52138 MW; 9043879ACE3E9B8 CRC64;

Query Match
Best Local Similarity 30.5%; Score 482; DB 1; Length 508;
Matches 123; Conservative 33; Mismatches 82; Indels 124; Gaps 11;

QY 39 PGQGPGRPGP-----PGAPGSGMGMGPFPGKQDQ-----DGHGDRGDSGEGPPGR 86
DB 146 PGERGPPGPPGAEKGLPGVPEKGDVGLMGPPGLDMRGATGLEDKDKDQDGTMPGA 205
QY 87 TGNNGKDGPK-----GKAGAIGRAPRGKGV 113
DB 206 PGLIGKSGPKDLPKPKBKRTGLPKLKGDLGERKPGMNGTQGEKDKDLGKIGAPSGSL 265
QY 114 NGTPGKHGTGPK-----KGPCKGEGPGLPGP-----CSC 143
DB 266 TGPWGONGQKEMGECEPTGEGKEGEGALPGPPGRLVGTGPGVGTNGTLPGVGLRQGL 325
QY 144 GS-----GHT-----KSAFSAV--T 157
DB 326 GSPGCKGAGRGPPGLRGMGPKGKPKPRGVGRGPKGQGTAGQIRAFSGVLPSS 385
QY 158 KSYPRERLPPIKFDKILNNEGHNASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGQ 217
DB 386 KSFPPGPGPVKFDILVNEEHMDPMLSKFNCTHPGVVFYHITVNRPLRALVINGV 445
QY 218 YRITFPDANTGNH-DVNSGSTTLALKQGDVWLQIFTSQNGLYDPYWTDSLTTGLIY 276
DB 446 KKRTTRSLYQGDIDQASNLALRLASGDQVWLETL-RDMNGV-YSSSESDSTFTGFLY 503
QY 277 AD 278
DB 504 AD 505

RESULT 15
COL-LEPMA STANDARD; PRT; 419 AA.
ID COL-LEPMA
AC P96085; Q91080;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DE Inner ear-specific collagen precursor (saccular collagen).
OS Lepomis macrochirus (Bluegill).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Perciformes; Percoidae;
OC Centrarchidae; Lepomis.
NCBI_TaxID=11106;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95167486; PubMed=7863311;
RA Davis J.G., Oberholzer J.C., Burns F.R., Greene M.I.;
RT "Molecular cloning and characterization of an inner ear-specific
RT structural protein.";
RL Science 267:1031-1034 (1995).
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Gibson T.;

```

```

RL Submitted (MAR-1995) to Swiss-Prot.
CC -1- FUNCTION: Forms a microstructural matrix within the otolithic
CC membrane (Probable).
CC -1- TISSUE SPECIFICITY: Specialized secretory supporting cells at the
CC outer perimeter of the saccular epithelium.
CC -1- SIMILARITY: Contains 1 C1q domain.
CC -1- CAUTION: This is a conceptual translation; a frameshift had to be
CC introduced for positions 391-419 so as to maximize the similarity
CC with other short-chain collagens.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U17431; AAA6978.1; ALT_FRAME.
DR HSSP; Q60994; 1C28.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR008161; C1q_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; C1q; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 2.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
DR Collagen; Extracellular matrix; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 419
FT DOMAIN 20 57
FT DOMAIN 58 274
FT DOMAIN 275 419
FT CARBOHYD 37 37
FT CARBOHYD 320 320
SQ SEQUENCE 419 AA; 43634 MW; 570CDB9675FC0F39 CRC64;

Query Match
Best Local Similarity 28.9%; Score 457.5; DB 1; Length 419;
Matches 114; Conservative 30; Mismatches 96; Indels 89; Gaps 9;

QY 40 GPGPPGPPGAP-----GPGSGMGMGPFPGKQDQDGHGDRGDSG--E 81
DB 88 GLRGPPGPPGALPGANGLNGDIGERKDGQVGLPGVPEIPKPKBKDPLGKDKGERGFS 147
QY 82 GPPGRTGNRGKPPGPKKAGAIGRAPRG-----KGVNTPGK 119
DB 148 GLMGDPPEBERGPPGLNGTKSIRGEGPWPGLAGTKLKGEGKLQGEKGERGPPGL 207
QY 120 HGTPEKKKPGKKGEPGLPPPCSGSGHT----- 148
DB 208 RGEWGLNGTGVVGEREPPLG-GKGDYGARPPGPPGGRGMAGLRGEKGLGVGRPG 266
QY 149 -----KSAFSAV--TKSYPRERLPPIKFDKILNNEGHNASSGKFCVCGVGI 194
DB 267 PKGPPGSGVQIRSAFSGVLPSSRSFPPEPLPVKFDVFNNGGHDPTLANKRVYTPGV 326
QY 195 YFTYDITLANKHLAIGLVHNGQYRITFPDANTGNH-DVNSGSTTLALKQGDVWLQIFY 253
DB 327 YLPSYHITVNRPLRALVINGVRKLRTRSLYQGDIDQASNLALHLTGDQVWLETL- 385
QY 254 SEQNGLYDPYWTDSLTTGLIYADDDP 282
DB 386 RDMNG-XYSSSESDSTFTSGFLYPTTKP 413

```

Search completed: December 25, 2004, 08:38:16  
Job time : 434.991 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 25, 2004, 08:12:38 / Search time 123.56 Seconds

(Without alignments)  
152,967 Million cell updates/sec

Title: US-10-621-787-2

Perfect score: 1582  
Sequence: 1 MIPWVLACALPCADPLG.....TDSIFGFLIYADDDDENV 285Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/PCFUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1582	100.0	285	4	US-09-552-204A-2
2	1487	94.0	294	3	US-09-188-930-294
3	1487	94.0	294	4	US-09-312-283C-294
4	1484	93.8	285	4	US-09-312-283C-382
5	981	62.0	198	3	US-09-188-930-138
6	981	62.0	198	4	US-09-312-283C-138
7	556.5	35.2	120	4	US-09-552-204A-12
8	513.5	32.5	162	3	US-09-227-357-436
9	431.5	27.3	244	2	US-08-463-911-7
10	431.5	27.3	244	2	US-09-140-804-3
11	431.5	27.3	244	4	US-09-336-536-20
12	431.5	27.3	244	4	US-09-530-423-1
13	431.5	27.3	244	4	US-09-686-838B-3
14	431.5	27.3	244	4	US-09-911-176B-48
15	431.5	27.3	244	4	US-09-552-225A-3
16	431.5	27.3	244	4	US-09-619-740-51
17	431.5	27.3	244	4	US-09-776-976-6
18	431.5	27.3	244	4	US-09-909-547-6
19	431.5	27.3	244	4	US-09-552-204A-3
20	425	26.9	247	2	US-08-463-911-2
21	425	26.9	247	4	US-09-776-976-4
22	425	26.9	247	4	US-09-909-547-4
23	424.5	26.8	244	4	US-09-568-852B-6
24	420.5	26.6	231	4	US-09-530-423-2
25	420	26.5	423	1	US-08-383-744-2
26	420	26.5	423	2	US-08-999-336-2
27	420	26.5	423	5	PCT-US96-01427-2

28	417	26.4	247	3	US-09-140-804-8	Sequence 8, Appli
29	417	26.4	247	3	US-09-118-408-3	Sequence 3, Appli
30	417	26.4	247	4	US-09-506-885-3	Sequence 8, Appli
31	417	26.4	247	4	US-09-686-838B-8	Sequence 3, Appli
32	417	26.4	247	4	US-09-911-176B-3	Sequence 3, Appli
33	417	26.4	247	4	US-09-619-740-3	Sequence 3, Appli
34	417	26.4	247	4	US-09-506-882-3	Sequence 3, Appli
35	415	26.2	243	3	US-09-188-930-295	Sequence 25, App
36	415	26.2	243	4	US-09-312-283C-295	Sequence 25, App
37	410	25.9	247	4	US-09-776-976-2	Sequence 2, Appli
38	410	25.9	247	4	US-09-909-547-2	Sequence 2, Appli
39	408	25.8	243	4	US-09-336-536-10	Sequence 10, Appli
40	403.5	25.5	228	4	US-09-336-536-11	Sequence 11, Appli
41	403	25.5	243	3	US-09-140-804-2	Sequence 2, Appli
42	403	25.5	243	4	US-09-336-536-2	Sequence 2, Appli
43	403	25.5	243	4	US-09-686-838B-2	Sequence 3, Appli
44	403	25.5	243	4	US-09-866-028-42	Sequence 42, Appli
45	403	25.5	243	4	US-10-140-002-362	Sequence 362, App
46	403	25.5	243	4	US-09-944-457-42	Sequence 42, Appli
47	398.5	25.2	228	4	US-09-336-536-4	Sequence 4, Appli
48	393	24.8	246	2	US-08-463-911-4	Sequence 4, Appli
49	366	23.1	245	4	US-09-552-225A-4	Sequence 4, Appli
50	366	23.1	245	4	US-09-552-204A-4	Sequence 4, Appli
51	355.5	22.5	258	4	US-09-976-594-815	Sequence 815, App
52	349.5	22.1	245	4	US-09-800-729-116	Sequence 116, App
53	348.5	22.0	245	3	US-09-140-804-4	Sequence 4, Appli
54	348.5	22.0	245	4	US-09-686-838B-4	Sequence 4, Appli
55	348.5	22.0	245	4	US-09-911-176B-49	Sequence 49, Appli
56	348.5	22.0	245	4	US-09-619-740-52	Sequence 52, Appli
57	348.5	22.0	245	4	US-09-800-729-85	Sequence 85, Appli
58	348.5	22.0	245	4	US-09-800-729-115	Sequence 115, App
59	348.5	22.0	245	4	US-09-800-729-118	Sequence 118, App
60	348.5	22.0	245	4	US-09-800-729-119	Sequence 119, App
61	345.5	21.8	245	4	US-09-311-021-104	Sequence 104, App
62	324	20.5	229	4	US-09-800-729-117	Sequence 117, App
63	311.5	19.7	1057	3	US-08-931-820-1	Sequence 1, Appli
64	311.5	19.7	1461	4	US-09-585-887-9	Sequence 9, Appli
65	311.5	19.7	1461	4	US-09-289-578-9	Sequence 9, Appli
66	311.5	19.7	1464	4	US-09-331-347C-21	Sequence 21, Appli
67	306	19.3	215	3	US-09-140-804-5	Sequence 5, Appli
68	306	19.3	215	4	US-09-686-838B-5	Sequence 5, Appli
69	306	19.3	215	4	US-09-911-176B-50	Sequence 50, Appli
70	306	19.3	215	4	US-09-619-740-53	Sequence 53, Appli
71	305.5	19.3	246	4	US-09-552-225A-12	Sequence 12, Appli
72	305	19.3	1341	3	US-08-963-825-18	Sequence 18, Appli
73	305	19.3	1341	3	US-09-500-811-18	Sequence 18, Appli
74	305	19.3	1341	3	US-09-570-573-18	Sequence 18, Appli
75	305	19.3	1341	3	US-09-548-608-18	Sequence 18, Appli
76	304	19.2	246	4	US-09-552-225A-2	Sequence 2, Appli
77	304	19.2	319	4	US-10-012-605C-2	Sequence 2, Appli
78	303	19.2	492	4	US-08-468-996-12	Sequence 12, Appli
79	302	19.1	310	3	US-09-219-849-47	Sequence 47, Appli
80	302	19.1	595	3	US-09-219-849-48	Sequence 48, Appli
81	302	19.1	595	3	US-09-219-849-50	Sequence 50, Appli
82	302	19.1	822	3	US-09-219-849-49	Sequence 49, Appli
83	293	18.5	1064	3	US-08-642-255-62	Sequence 62, Appli
84	288.5	18.2	1078	3	US-08-963-825-21	Sequence 21, Appli
85	288.5	18.2	1078	3	US-09-500-811-21	Sequence 21, Appli
86	288.5	18.2	1078	3	US-09-570-573-21	Sequence 21, Appli
87	288.5	18.2	1078	3	US-09-548-608-21	Sequence 21, Appli
88	288	18.2	1330	1	US-08-642-255-12	Sequence 12, Appli
89	288	18.2	408	1	US-07-609-716-65	Sequence 65, Appli
90	288	18.2	408	3	US-08-475-411A-65	Sequence 65, Appli
91	288	18.2	408	3	US-08-478-029A-65	Sequence 65, Appli
92	286.5	18.1	1057	3	US-08-931-820-4	Sequence 4, Appli
93	286	18.1	1418	3	US-08-963-825-20	Sequence 20, Appli
94	286	18.1	1418	3	US-09-010-999-1	Sequence 1, Appli
95	286	18.1	1418	3	US-09-500-811-20	Sequence 20, Appli
96	286	18.1	1418	3	US-09-570-573-20	Sequence 20, Appli
97	286	18.1	1418	3	US-09-548-608-20	Sequence 20, Appli
98	285.5	18.0	252	1	US-08-642-255-61	Sequence 61, Appli
99	285	18.0	357	1	US-07-609-716-66	Sequence 66, Appli
100	285	18.0	357	1	US-08-642-255-33	Sequence 33, Appli

## ALIGNMENTS

RESULT 1  
US-09-552-204A-2  
Sequence 2, Application US/09552204A  
Patent No. 6620909  
GENERAL INFORMATION:  
APPLICANT: Piddington, Christopher S.  
APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGY ZACR2  
FILE REFERENCE: 99-08  
CURRENT APPLICATION NUMBER: US/09/552,204A  
CURRENT FILING DATE: 2000-04-19  
PRIOR APPLICATION NUMBER: 60/130,207  
PRIOR FILING DATE: 1999-04-20  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 285  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-552-204A-2

Query Match 100.0%; Score 1582; DB 4; Length 285;  
Best Local Similarity 100.0%; Pred. No. 2,4e-132;  
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPFGPAGAPSGMMGRM 60  
DB 1 MIPWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPFGPAGAPSGMMGRM 60  
QY 61 GPPKQDQDGDHGRGSGEGPGRGTGNRGKPKPKAKAIGAGRGPGVNGTGK 120  
DB 61 GPPKQDQDGDHGRGSGEGPGRGTGNRGKPKPKAKAIGAGRGPGVNGTGK 120  
QY 121 GTPKPKPKKKGKPGGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 180  
DB 121 GTPKPKPKKKGKPGGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 180  
QY 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHVA 240  
DB 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHVA 240  
QY 241 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285  
DB 241 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285

RESULT 2  
US-09-188-930-294  
Sequence 294, Application US/09188930A  
Patent No. 6150502  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murlison, James Greg  
TITLE OF INVENTION: Compositions Isolated from Skin Cells  
FILE REFERENCE: 11000,1011c1  
CURRENT APPLICATION NUMBER: US/09/188,930A  
CURRENT FILING DATE: 1998-11-09  
NUMBER OF SEQ ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 294  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Rat  
US-09-188-930-294

Query Match 94.0%; Score 1487; DB 3; Length 294;  
Best Local Similarity 94.0%; Pred. No. 6,4e-124;  
Matches 268; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 MIPWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPFGPAGAPSGMMGRM 60  
DB 10 MISWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPFGPAGAPSGMMGRM 69  
QY 61 GPPKQDQDGDHGRGSGEGPGRGTGNRGKPKPKAKAIGAGRGPGVNGTGK 120  
DB 70 GPPKQDQDGDHGRGSGEGPGRGTGNRGKPKPKAKAIGAGRGPGVNGTGK 129  
QY 121 GTPKPKPKKKGKPGGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 180  
DB 130 GTPKPKPKKKGKPGGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 189  
QY 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHVA 240  
DB 190 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHVA 249  
QY 241 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285  
DB 250 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 294

RESULT 3  
US-09-312-283C-294  
Sequence 294, Application US/09312283C  
Patent No. 6573095  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murlison, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions Isolated from Skin Cells  
FILE REFERENCE: 11000,1011c2  
CURRENT APPLICATION NUMBER: US/09/312,283C  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 425  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 294  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Mouse  
US-09-312-283C-294

Query Match 94.0%; Score 1487; DB 4; Length 294;  
Best Local Similarity 94.0%; Pred. No. 6,4e-124;  
Matches 268; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 MIPWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPFGPAGAPSGMMGRM 60  
DB 10 MISWVLLAALPCADPLGAFARDFRKGSPOLVCSLPQGPFGPAGAPSGMMGRM 69  
QY 61 GPPKQDQDGDHGRGSGEGPGRGTGNRGKPKPKAKAIGAGRGPGVNGTGK 120  
DB 70 GPPKQDQDGDHGRGSGEGPGRGTGNRGKPKPKAKAIGAGRGPGVNGTGK 129  
QY 121 GTPKPKPKKKGKPGGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 180  
DB 130 GTPKPKPKKKGKPGGLPGPCSCSGHTKSAFSAVATKSYPRRLPIKFDKILNNEG 189  
QY 181 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHVA 240  
DB 190 NASSGKFCVCGVPGIYYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHVA 249  
QY 241 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285  
DB 250 LKQGDVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 294



```
RESULT 4
US-09-312-283C-382
; Sequence 382, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 382
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-382

Query Match          93.8%; Score 1484; DB 4; Length 285;
Best Local Similarity 93.7%; Pred. No. 1,1e-123;
Matches 267; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCAADPMLGAFARDDPRKSPQLVCSLPFGQPPGPPGAGPBGMMGRM 60
D 1 MISWMLLACALPCAADPMLGAFARDDPRKSPQLVCSLPFGQPPGPPGAGPBGMMGRM 60
QY 61 GFPKQDQDGDHGRGDSGSEBPPGRTGNRKGPPKAGAGIAGRAGRPKGVNGTPEKH 120
D 61 GFPKQDQDGDHGRGDSGSEBPPGRTGNRKGPPKAGAGIAGRAGRPKGVNGTPEKH 120
QY 121 GTPKPKPKKGGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 180
D 121 GTPKPKPKKGGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 180
QY 181 NASGKRFVCVCGIYFTYDITLANKHLAIGLVHNGVORITPTANTGNHVAAGSTTLA 240
D 181 NASGKRFVCVCGIYFTYDITLANKHLAIGLVHNGVORITPTANTGNHVAAGSTTLA 240
QY 241 LKQDEVWLQIFYSQNGLFTYDPTWDSLFTGFLIYADQDDPNEY 285
D 241 LKQDEVWLQIFYSQNGLFTYDPTWDSLFTGFLIYADQDDPNEY 285

RESULT 5
US-09-188-930-138
; Sequence 138, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 138
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Rat
US-09-188-930-138
```

```
Query Match          62.0%; Score 981; DB 3; Length 198;
Best Local Similarity 92.6%; Pred. No. 2.6e-79;
Matches 175; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCAADPMLGAFARDDPRKSPQLVCSLPFGQPPGPPGAGPBGMMGRM 60
D 10 MISWMLLACALPCAADPMLGAFARDDPRKSPQLVCSLPFGQPPGPPGAGPBGMMGRM 69
QY 61 GFPKQDQDGDHGRGDSGSEBPPGRTGNRKGPPKAGAGIAGRAGRPKGVNGTPEKH 120
D 70 GFPKQDQDGDHGRGDSGSEBPPGRTGNRKGPPKAGAGIAGRAGRPKGVNGTPEKH 129
QY 121 GTPKPKPKKGGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 180
D 130 GTPKPKPKKGGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 189
QY 181 NASGKRFVC 189
D 190 NASGKRFVC 198

RESULT 6
US-09-312-283C-138
; Sequence 138, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-138

Query Match          62.0%; Score 981; DB 4; Length 198;
Best Local Similarity 92.6%; Pred. No. 2.6e-79;
Matches 175; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MIPVLLACALPCAADPMLGAFARDDPRKSPQLVCSLPFGQPPGPPGAGPBGMMGRM 60
D 10 MISWMLLACALPCAADPMLGAFARDDPRKSPQLVCSLPFGQPPGPPGAGPBGMMGRM 69
QY 61 GFPKQDQDGDHGRGDSGSEBPPGRTGNRKGPPKAGAGIAGRAGRPKGVNGTPEKH 120
D 70 GFPKQDQDGDHGRGDSGSEBPPGRTGNRKGPPKAGAGIAGRAGRPKGVNGTPEKH 129
QY 121 GTPKPKPKKGGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 180
D 130 GTPKPKPKKGGKGGPGLPGPCSCSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGHHY 189
QY 181 NASGKRFVC 189
D 190 NASGKRFVC 198

RESULT 7
US-09-552-204A-12
; Sequence 12, Application US/09552204A
; Patent No. 6620909
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
```

APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACRBP2  
FILE REFERENCE: 99-08  
CURRENT APPLICATION NUMBER: US/09/552,204A  
CURRENT FILING DATE: 2000-04-19  
PRIOR APPLICATION NUMBER: 60/130,207  
PRIOR FILING DATE: 1999-04-20  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 120  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)-(120)  
OTHER INFORMATION: Bach Xaa is independently any amino acid.  
US-09-552-204A-12

Query Match 35.2%; Score 556.5; DB 4; Length 120;  
Best Local Similarity 88.4%; Pred. No. 5.6e-42;  
Matches 107; Conservative 2; Mismatches 9; Indels 3; Gaps 2;

QY 167 IKPKILMNEGGHYNASSGKFCVCGPYGYTYDITL--ANKHLAIGLVNNGQYRIRTFD 224  
DB 1 IKPKILMNEGGHYNASSGKFCVCGPYGYTYDITL--ANKHLAIGLVNNGQYRIRTFD 59  
QY 225 ANTNHNVASGSTITALAKQGEVWLQIPIYSBONGLFYDPTWDSLFTGFLLYADODDPNE 284  
DB 60 ANTNHNVASGSTITALAKQGEVWLQIPIYSBONGLFYDPTWDSLFTGFLLYADODDPNE 119  
QY 285 V 285  
DB 120 V 120

RESULT 8  
US-09-227-357-436  
Sequence 436, Application US/09227357  
Patent No. 6342581  
GENERAL INFORMATION:  
APPLICANT: Fischer et al.  
TITLE OF INVENTION: 123 Human Secreted Proteins  
FILE REFERENCE: P2010P1  
CURRENT APPLICATION NUMBER: US/09/227,357  
CURRENT FILING DATE: 1999-01-08  
EARLIER APPLICATION NUMBER: PCT/US98/13684  
EARLIER FILING DATE: 1998-07-07  
EARLIER APPLICATION NUMBER: 60/051,926  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,793  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,925  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,929  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,803  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,732  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,931  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,932  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,916  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,930  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,918  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,920  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,733

EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,795  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,919  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,928  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/055,722  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,723  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,948  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,949  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,953  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,950  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,947  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,964  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/056,360  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,684  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,984  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,954  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/058,785  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,664  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,660  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,661  
EARLIER FILING DATE: 1997-09-12  
NUMBER OF SEQ ID NOS: 672  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 436  
LENGTH: 162  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (1)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (33)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (48)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-227-357-436

Query Match 32.5%; Score 513.5; DB 3; Length 162;  
Best Local Similarity 65.1%; Pred. No. 5.3e-38;  
Matches 99; Conservative 17; Mismatches 27; Indels 9; Gaps 2;

QY 134 BPGLPBCSGSGHTKSAFSAVAATKYSRRLPIKPKILMNEGG-----HYNASGKFCV 188  
DB 4 DPGLPBCSGSIVLKSASFVSGITTSYPERKRLPIPKKVLIPKXALQPCHRSSS----- 59  
QY 189 CGVPGIYFTYDITLANKHLAIGLVNNGQYRIRTPDANTGNHNVASGSTITALAKQGEVW 248  
DB 60 VLSGCIYFSYDITLANKHLAIGLVNNGQYRIRTPDANTGNHNVASGSTITYIYLPEDVW 119  
QY 249 LQIYSBONGLFYDPTWDSLFTGFLLYADOD 280

Db 120 LEIFTDQNGLFSDPGMADSLFSGFLTYVTD 151

RESULT 9

US-08-463-911-7

Sequence 7, Application US/08463911

Patent No. 5869330

GENERAL INFORMATION:

APPLICANT: Scherer, Philipp E.

APPLICANT: Lodish, Harvey F.

TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED

TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Millitia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,911

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: WH195-05

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 244 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-463-911-7

Query Match 27.3%; Score 431.5; DB 2; Length 244;

Best Local Similarity 40.2%; Pred. No. 1.6e-30;

Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPQLVCSLPG-----PGGP-----PGPPGAPGSPGMMGMPGPKDQGDGDDRDGDSG 79

Db 5 GAVLLALALPGHDDQETTTQGPVLLPLPKGA--CTGMAA--GIPGHPGHN-----50

Qy 80 EEGPPGRTGNRGKGPCKAGATGRAGPRGPGVNGTPGKHGTGPKKGPCKGKGPGLPG 139

Db 51 --GAPGRDGRDGTGEGKEKDPGLIGPKDIDGTVGAPGAPRGPGLIGRKKEPGE--106

Qy 140 PCSCSGHTSASFVAATYKSPRERLPIKFDKILMNEGSHYNASSGKFCVCGVPIYFTY 199

Db 107 ----GAVYRSAPFVGLETYYTINMPIRFTKIFYNOQNHDSGTGKFKHCNIPGLYYFAY 162

Qy 200 DITLANHGLAIGLVHNGQYRIRTFDA--NTGNHDVAVSGSTIALAKQSGEVLQIF--YSEON 257

Db 163 HITVYMDVKVSLPKKOKAMLFITYDOYQENNVDAAGSVLLHLHVGQVWLQVYGEGERN 222

Qy 258 GLFYDPYWTDSLFTGFLIYAD 278

Db 223 GLYADND--NDSTFTGFLIYHD 242

RESULT 10

US-09-140-804-3

Sequence 3, Application US/09140804

Patent No. 6197930

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Humes, Jacqueline M.

TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS

FILE REFERENCE: 97-49

CURRENT APPLICATION NUMBER: US/09/140,804

CURRENT FILING DATE: 1998-08-26

EARLIER APPLICATION NUMBER: 60/056,983

EARLIER FILING DATE: 1997-08-26

NUMBER OF SEQ ID NOS: 47

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 3

LENGTH: 244

TYPE: PRT

ORGANISM: Homo sapiens

US-09-140-804-3

Query Match 27.3%; Score 431.5; DB 3; Length 244;

Best Local Similarity 40.2%; Pred. No. 1.6e-30;

Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPQLVCSLPG-----PGGP-----PGPPGAPGSPGMMGMPGPKDQGDGDDRDGDSG 79

Db 5 GAVLLALALPGHDDQETTTQGPVLLPLPKGA--CTGMAA--GIPGHPGHN-----50

Qy 80 EEGPPGRTGNRGKGPCKAGATGRAGPRGPGVNGTPGKHGTGPKKGPCKGKGPGLPG 139

Db 51 --GAPGRDGRDGTGEGKEKDPGLIGPKDIDGTVGAPGAPRGPGLIGRKKEPGE--106

Qy 140 PCSCSGHTSASFVAATYKSPRERLPIKFDKILMNEGSHYNASSGKFCVCGVPIYFTY 199

Db 107 ----GAVYRSAPFVGLETYYTINMPIRFTKIFYNOQNHDSGTGKFKHCNIPGLYYFAY 162

Qy 200 DITLANHGLAIGLVHNGQYRIRTFDA--NTGNHDVAVSGSTIALAKQSGEVLQIF--YSEON 257

Db 163 HITVYMDVKVSLPKKOKAMLFITYDOYQENNVDAAGSVLLHLHVGQVWLQVYGEGERN 222

Qy 258 GLFYDPYWTDSLFTGFLIYAD 278

Db 223 GLYADND--NDSTFTGFLIYHD 242

RESULT 11

US-09-336-536-20

Sequence 20, Application US/09336536

Patent No. 6406884

GENERAL INFORMATION:

APPLICANT: Leiby, K.

APPLICANT: McKay, C.

APPLICANT: Boesone, S.

TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF

FILE REFERENCE: 7853-144

CURRENT APPLICATION NUMBER: US/09/336,536

CURRENT FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 75

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 20

LENGTH: 244

TYPE: PRT

ORGANISM: Homo sapiens

US-09-336-536-20

Query Match 27.3%; Score 431.5; DB 4; Length 244;

Best Local Similarity 40.2%; Pred. No. 1.6e-30;

Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPQLVCSLPG-----PGGP-----PGPPGAPGSPGMMGMPGPKDQGDGDDRDGDSG 79

Db 5 GAVLLALALPGHDDQETTTQGPVLLPLPKGA--CTGMAA--GIPGHPGHN-----50

Qy 80 EEGPPGRTGNRGKGPCKAGATGRAGPRGPGVNGTPGKHGTGPKKGPCKGKGPGLPG 139

Db 51 --GAPGRDGRDGTGEGKEKDPGLIGPKDIDGTVGAPGAPRGPGLIGRKKEPGE--106

Qy 140 PCSCGSHGTSAPSAVATKSYPRERLPKPKDKILMEGSHYNASSGKFCVGPQIYYFTY 199  
 Db 107 ----GAYYRSASFVSGLETYYTIPNMPIRFTKIFYNQNMHYDSTGKFKHCNIPGLYYFAY 162  
 Qy 200 DITLANHGLAIGLVHNGQYRIRTPDA-NTGNHVAASGSTITALKQGEVWLQIF-YSEON 257  
 Db 163 HITVYMDVKVSLFKKOKAMLFYDQYQENNVDQASGSVLLHLFVGDQVWLQVYVGEGERN 222  
 Qy 258 GLFYDPYWTDSLFTGFLLIYAD 278  
 Db 223 GLYADND-NDSTFTGFLLIYHD 242

## RESULT 12

US-09-530-423-1  
 ; Sequence 1, Application US/09530423  
 ; Patent No. 6461821  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Otsuka Pharmaceutical Co., Ltd.  
 ; TITLE OF INVENTION: Smooth muscle growth inhibitory composition, a  
 ; TITLE OF INVENTION: diagnostic method for arteriosclerosis and a kit  
 ; TITLE OF INVENTION: cherefor  
 ; FILE REFERENCE: P98-51  
 ; CURRENT APPLICATION NUMBER: US/09/530,423  
 ; CURRENT FILING DATE: 2000-05-01  
 ; PRIOR APPLICATION NUMBER: JP H9-297569  
 ; PRIOR FILING DATE: 1997-10-29  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 244  
 ; TYPE: PRT  
 ; ORGANISM: Abdominal fat tissue from myoma uteri  
 US-09-530-423-1

Query Match 27.3%; Score 431.5; DB 4; Length 244;  
 Best Local Similarity 40.2%; Pred. No. 1.6e-30;  
 Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPOLVCSLPG-----PQGP----PGPPAPGSGMGMGPFPGKQGDGHDGRDSDG 79  
 Db 5 GAVLLALPQHDOETTTQGPVLLPLPKGA--CTGMA--GIPGPHGN----- 50  
 Qy 80 EEPGRTGNRKGPCKAGAIGRAPRGKGVNTPGKHGTGPKKGPCKGPGPLPG 139  
 Db 51 --GAPGDRDGTGPEGEKGDPLIGPKDIGETVPGAEGRGPGIQRKGPGE-- 106  
 Qy 140 PCSCGSHGTSAPSAVATKSYPRERLPKPKDKILMEGSHYNASSGKFCVGPQIYYFTY 199  
 Db 107 ----GAYYRSASFVSGLETYYTIPNMPIRFTKIFYNQNMHYDSTGKFKHCNIPGLYYFAY 162  
 Qy 200 DITLANHGLAIGLVHNGQYRIRTPDA-NTGNHVAASGSTITALKQGEVWLQIF-YSEON 257  
 Db 163 HITVYMDVKVSLFKKOKAMLFYDQYQENNVDQASGSVLLHLFVGDQVWLQVYVGEGERN 222  
 Qy 258 GLFYDPYWTDSLFTGFLLIYAD 278  
 Db 223 GLYADND-NDSTFTGFLLIYHD 242

## RESULT 13

US-09-686-838B-3  
 ; Sequence 3, Application US/09686838B  
 ; Patent No. 6482612  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sheppard, Paul O.  
 ; TITLE OF INVENTION: Adipocyte-Specific Protein Homologs  
 ; FILE REFERENCE: 97-49D1  
 ; CURRENT APPLICATION NUMBER: US/09/686,838B  
 ; CURRENT FILING DATE: 2000-10-10  
 ; PRIOR APPLICATION NUMBER: US 09/140,804

; PRIOR FILING DATE: 1998-08-26  
 ; PRIOR APPLICATION NUMBER: US 60/056,983  
 ; PRIOR FILING DATE: 1997-08-26  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 244  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-686-838B-3

Query Match 27.3%; Score 431.5; DB 4; Length 244;  
 Best Local Similarity 40.2%; Pred. No. 1.6e-30;  
 Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPOLVCSLPG-----PQGP----PGPPAPGSGMGMGPFPGKQGDGHDGRDSDG 79  
 Db 5 GAVLLALPQHDOETTTQGPVLLPLPKGA--CTGMA--GIPGPHGN----- 50  
 Qy 80 EEPGRTGNRKGPCKAGAIGRAPRGKGVNTPGKHGTGPKKGPCKGPGPLPG 139  
 Db 51 --GAPGDRDGTGPEGEKGDPLIGPKDIGETVPGAEGRGPGIQRKGPGE-- 106  
 Qy 140 PCSCGSHGTSAPSAVATKSYPRERLPKPKDKILMEGSHYNASSGKFCVGPQIYYFTY 199  
 Db 107 ----GAYYRSASFVSGLETYYTIPNMPIRFTKIFYNQNMHYDSTGKFKHCNIPGLYYFAY 162  
 Qy 200 DITLANHGLAIGLVHNGQYRIRTPDA-NTGNHVAASGSTITALKQGEVWLQIF-YSEON 257  
 Db 163 HITVYMDVKVSLFKKOKAMLFYDQYQENNVDQASGSVLLHLFVGDQVWLQVYVGEGERN 222  
 Qy 258 GLFYDPYWTDSLFTGFLLIYAD 278  
 Db 223 GLYADND-NDSTFTGFLLIYHD 242

## RESULT 14

US-09-911-176B-48  
 ; Sequence 48, Application US/09911176B  
 ; Patent No. 6516403  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sheppard, Paul O.  
 ; TITLE OF INVENTION: ANTIBODIES THAT BIND AN  
 ; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG  
 ; FILE REFERENCE: 97-30D1  
 ; CURRENT APPLICATION NUMBER: US/09/911,176B  
 ; CURRENT FILING DATE: 2001-07-23  
 ; PRIOR APPLICATION NUMBER: 09/118,408  
 ; PRIOR FILING DATE: 1998-07-17  
 ; PRIOR APPLICATION NUMBER: 60/053,154  
 ; PRIOR FILING DATE: 1997-07-18  
 ; NUMBER OF SEQ ID NOS: 52  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 48  
 ; LENGTH: 244  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-911-176B-48

Query Match 27.3%; Score 431.5; DB 4; Length 244;  
 Best Local Similarity 40.2%; Pred. No. 1.6e-30;  
 Matches 105; Conservative 33; Mismatches 88; Indels 35; Gaps 9;

Qy 30 GSPOLVCSLPG-----PQGP----PGPPAPGSGMGMGPFPGKQGDGHDGRDSDG 79  
 Db 5 GAVLLALPQHDOETTTQGPVLLPLPKGA--CTGMA--GIPGPHGN----- 50  
 Qy 80 EEPGRTGNRKGPCKAGAIGRAPRGKGVNTPGKHGTGPKKGPCKGPGPLPG 139  
 Db 51 --GAPGDRDGTGPEGEKGDPLIGPKDIGETVPGAEGRGPGIQRKGPGE-- 106  
 Qy 140 PCSCGSHGTSAPSAVATKSYPRERLPKPKDKILMEGSHYNASSGKFCVGPQIYYFTY 199  
 Db 163 HITVYMDVKVSLFKKOKAMLFYDQYQENNVDQASGSVLLHLFVGDQVWLQVYVGEGERN 222



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:38:38 ; Search time 501.456 Seconds  
(without alignments)  
204.102 Million cell updates/sec

Title: US-10-621-787-2

Perfect score: 1582

Sequence: 1 MIMWVLLACMLPCADPLG.....TDSLPTGFLIYADDDNEV 265

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1595201 seqs, 359116952 residues

Total number of hits satisfying chosen parameters: 1595201

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*

9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/2/pubppaa/US09B\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*

12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*

13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/2/pubppaa/US10B\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/2/pubppaa/US10D\_PUBCOMB.pep:\*

17: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*

18: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*

19: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*

20: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1582	100.0	285 14 US-10-234-000-5	Sequence 5, Appl1
2	1582	100.0	285 14 US-10-411-120-60	Sequence 60, Appl1
3	1582	100.0	285 14 US-10-411-120-88	Sequence 88, Appl1
4	1582	100.0	285 15 US-10-621-787-2	Sequence 2, Appl1
5	1582	100.0	314 15 US-10-264-049-2619	Sequence 2619, Ap
6	1582	100.0	330 15 US-10-236-055A-6	Sequence 6, Appl1
7	1582	100.0	330 15 US-10-162-335-30	Sequence 30, Appl1
8	1578	99.7	331 15 US-10-220-120-354	Sequence 354, App
9	1492	94.3	294 14 US-10-236-055A-8	Sequence 8, Appl1
10	1487	94.0	294 10 US-09-866-050A-294	Sequence 294, App
11	1484	93.8	285 10 US-09-866-050A-382	Sequence 382, App
12	1154	72.9	217 9 US-09-770-906-2	Sequence 2, Appl1
13	981	62.0	198 10 US-09-866-050A-138	Sequence 138, App

14	898.5	56.8	289 14 US-10-234-000-15	Sequence 15, Appl1
15	898	56.8	289 10 US-09-866-050A-630	Sequence 630, App
16	895	56.6	273 17 US-10-758-846-60	Sequence 60, Appl1
17	895	56.6	289 14 US-10-411-120-59	Sequence 59, Appl1
18	895	56.6	289 15 US-10-451-168-106	Sequence 106, App
19	895	56.6	289 16 US-10-408-765A-1762	Sequence 1762, App
20	895	56.6	289 17 US-10-758-846-59	Sequence 59, Appl1
21	895	56.6	289 17 US-10-758-846-75	Sequence 75, Appl1
22	895	56.6	303 14 US-10-234-000-2	Sequence 2, Appl1
23	895	56.6	303 14 US-10-203-708-31	Sequence 31, Appl1
24	879.5	55.6	288 14 US-10-203-708-30	Sequence 30, Appl1
25	675	42.7	126 14 US-10-411-120-107	Sequence 107, App
26	671	42.4	125 17 US-10-758-846-101	Sequence 101, App
27	561.5	35.5	120 15 US-10-621-787-12	Sequence 12, Appl1
28	521.5	33.0	333 17 US-10-758-846-70	Sequence 70, Appl1
29	521.5	33.0	333 14 US-10-411-120-69	Sequence 69, Appl1
30	521.5	33.0	333 14 US-10-411-120-97	Sequence 97, Appl1
31	521.5	33.0	333 15 US-10-423-582-2	Sequence 2, Appl1
32	521.5	33.0	333 17 US-10-726-699-72	Sequence 72, Appl1
33	521.5	33.0	333 17 US-10-726-699-95	Sequence 95, Appl1
34	519.5	32.8	298 15 US-10-423-584-2	Sequence 2, Appl1
35	519.5	32.8	333 15 US-10-112-944-412	Sequence 412, App
36	519.5	32.8	333 17 US-10-758-846-10	Sequence 10, Appl1
37	519.5	32.8	680 14 US-10-177-293-59	Sequence 59, Appl1
38	519.5	32.8	680 15 US-10-058-270A-80	Sequence 80, Appl1
39	517.5	32.7	124 17 US-10-758-846-107	Sequence 107, App
40	516.5	32.6	680 14 US-10-295-027-110	Sequence 110, App
41	516.5	32.6	680 17 US-10-643-795A-118	Sequence 118, App
42	514.5	32.5	162 10 US-09-983-802-436	Sequence 436, App
43	513.5	32.5	162 10 US-09-984-490-436	Sequence 436, App
44	513.5	32.5	162 11 US-09-973-278-565	Sequence 565, App
45	513.5	32.5	162 14 US-10-411-120-101	Sequence 101, App
46	513.5	32.5	288 15 US-10-112-944-410	Sequence 410, App
47	507	32.0	288 17 US-10-758-846-4	Sequence 4, Appl1
48	506	32.0	269 17 US-10-758-846-5	Sequence 5, Appl1
49	482	30.5	508 17 US-10-758-846-90	Sequence 90, Appl1
50	479	30.3	98 9 US-09-770-906-4	Sequence 4, Appl1
51	466	29.1	482 17 US-10-758-846-92	Sequence 92, Appl1
52	460	29.1	493 17 US-10-758-846-89	Sequence 89, Appl1
53	457.5	28.9	459 16 US-10-408-765A-2554	Sequence 2554, App
54	457.5	28.9	459 17 US-10-758-846-28	Sequence 28, Appl1
55	457.5	28.9	477 17 US-10-758-846-27	Sequence 27, Appl1
56	457.5	28.9	477 17 US-10-758-846-91	Sequence 91, Appl1
57	456	28.8	481 17 US-10-758-846-93	Sequence 93, Appl1
58	453.5	28.7	481 17 US-10-758-846-88	Sequence 88, Appl1
59	451.5	28.5	438 17 US-10-758-846-78	Sequence 78, Appl1
60	450	28.4	438 15 US-10-138-588-42	Sequence 42, Appl1
61	446	28.2	703 14 US-10-219-449-4	Sequence 4, Appl1
62	444	28.1	717 14 US-10-219-449-2	Sequence 2, Appl1
63	444	28.1	733 15 US-10-138-588-40	Sequence 40, Appl1
64	443	28.0	638 13 US-10-001-887-108	Sequence 108, App
65	438.5	27.7	703 14 US-10-236-055A-32	Sequence 32, Appl1
66	436	27.6	744 14 US-10-411-120-65	Sequence 65, Appl1
67	436	27.6	744 14 US-10-301-822-39	Sequence 39, Appl1
68	436	27.6	744 17 US-10-758-846-25	Sequence 25, Appl1
69	436	27.6	717 17 US-10-758-846-70	Sequence 70, Appl1
70	434	27.4	744 14 US-10-236-055A-30	Sequence 30, Appl1
71	434	27.4	744 17 US-10-758-846-24	Sequence 24, Appl1
72	434	27.4	244 9 US-09-776-976-6	Sequence 6, Appl1
73	431.5	27.3	244 9 US-09-758-055-6	Sequence 6, Appl1
74	431.5	27.3	244 9 US-09-909-547-6	Sequence 6, Appl1
75	431.5	27.3	244 9 US-09-911-176B-48	Sequence 48, Appl1
76	431.5	27.3	244 14 US-10-180-762-51	Sequence 51, Appl1
77	431.5	27.3	244 14 US-10-231-814-6	Sequence 6, Appl1
78	431.5	27.3	244 14 US-10-234-000-4	Sequence 4, Appl1
79	431.5	27.3	244 14 US-10-360-186-51	Sequence 51, Appl1
80	431.5	27.3	244 14 US-10-321-164-3	Sequence 3, Appl1
81	431.5	27.3	244 14 US-10-189-493-1	Sequence 1, Appl1
82	431.5	27.3	244 14 US-10-197-293-3	Sequence 3, Appl1
83	431.5	27.3	244 14 US-10-325-717-1	Sequence 7, Appl1
84	431.5	27.3	244 14 US-10-411-120-67	Sequence 67, Appl1
85	431.5	27.3	244 15 US-10-379-747-6	Sequence 6, Appl1
86	431.5	27.3		

```
87 431.5 27.3 244 15 US-10-379-747-8 Sequence 8, Appli
88 431.5 27.3 244 15 US-10-379-747-10 Sequence 10, Appli
89 431.5 27.3 244 15 US-10-621-787-3 Sequence 10, Appli
90 431.5 27.3 244 15 US-10-285-833-6 Sequence 6, Appli
91 431.5 27.3 244 15 US-10-723-860-1048 Sequence 1048, Ap
92 431.5 27.3 250 17 US-10-379-747-12 Sequence 12, Appli
93 431 27.2 744 14 US-10-171-311-40 Sequence 40, Appli
94 431 27.2 744 14 US-10-301-822-37 Sequence 37, Appli
95 425 26.9 247 9 US-09-776-976-4 Sequence 4, Appli
96 425 26.9 247 9 US-09-758-055-4 Sequence 4, Appli
97 425 26.9 247 9 US-09-909-547-4 Sequence 4, Appli
98 425 26.9 247 14 US-10-231-814-4 Sequence 4, Appli
99 425 26.9 247 15 US-10-285-833-4 Sequence 4, Appli
100 424.5 26.8 244 14 US-10-376-460-6 Sequence 6, Appli
```

## ALIGNMENTS

```
RESULT 1
US-10-234-000-5
; Sequence 5, Application US/10234000
; Publication No. US20030129698A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN HOMOLOG ZACRP7
; FILE REFERENCE: 99-31
; CURRENT APPLICATION NUMBER: US/10/234,000
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US/09/577,298
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/136,289
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/145,589
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/158,448
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-234-000-5
```

```
Query Match 100.0%; Score 1582; DB 14; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.4e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPWVLLACALPCADADPLGAFARDRFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
DB 1 MIPWVLLACALPCADADPLGAFARDRFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
QY 61 GPFKDGQDHDGDRGDSGSEGGPRTGNRGKPGPKAKAIGRAGRGKGVNGTGGK 120
DB 61 GPFKDGQDHDGDRGDSGSEGGPRTGNRGKPGPKAKAIGRAGRGKGVNGTGGK 120
QY 121 GTGKKGPKKKGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGK 180
DB 121 GTGKKGPKKKGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGK 180
QY 181 NASGKFCVCGVPGLIYFTYDITLANKHLAIGLVNNGYRIRTPDANTGNHDVAGSTTLA 240
DB 181 NASGKFCVCGVPGLIYFTYDITLANKHLAIGLVNNGYRIRTPDANTGNHDVAGSTTLA 240
QY 241 LKQGDVWMLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285
DB 241 LKQGDVWMLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285
```

RESULT 2  
US-10-411-120-60

```
; Sequence 60, Application US/10411120
; Publication No. US20030224425A1
; GENERAL INFORMATION:
; APPLICANT: Blondel et al.
; TITLE OF INVENTION: ACRP-30-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P3131P1
; CURRENT APPLICATION NUMBER: US/10/411,120
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/328,419
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: PCT/US02/32432
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 60
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-120-60
```

```
Query Match 100.0%; Score 1582; DB 14; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.4e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPWVLLACALPCADADPLGAFARDRFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
DB 1 MIPWVLLACALPCADADPLGAFARDRFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
QY 61 GPFKDGQDHDGDRGDSGSEGGPRTGNRGKPGPKAKAIGRAGRGKGVNGTGGK 120
DB 61 GPFKDGQDHDGDRGDSGSEGGPRTGNRGKPGPKAKAIGRAGRGKGVNGTGGK 120
QY 121 GTGKKGPKKKGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGK 180
DB 121 GTGKKGPKKKGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGKPGK 180
QY 181 NASGKFCVCGVPGLIYFTYDITLANKHLAIGLVNNGYRIRTPDANTGNHDVAGSTTLA 240
DB 181 NASGKFCVCGVPGLIYFTYDITLANKHLAIGLVNNGYRIRTPDANTGNHDVAGSTTLA 240
QY 241 LKQGDVWMLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285
DB 241 LKQGDVWMLQIFYSEONGLFYDPYWTDSLFTGFLIYADODDPNEV 285
```

## RESULT 3

```
US-10-411-120-88
; Sequence 88, Application US/10411120
; Publication No. US20030224425A1
; GENERAL INFORMATION:
; APPLICANT: Blondel et al.
; TITLE OF INVENTION: ACRP-30-like Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P3131P1
; CURRENT APPLICATION NUMBER: US/10/411,120
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/328,419
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: PCT/US02/32432
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 88
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-120-88
```

```
Query Match 100.0%; Score 1582; DB 14; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.4e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPWVLLACALPCADADPLGAFARDRFRKSPOLVCSLPGPQGPBPFGAPGPGSGMMGRM 60
```

```
Db 1 MIPVLLAALPCADPLGAFARDFRKS POLVCSLPGPQPPGPPGAPGPGMMGRM 60
Qy 61 GPPKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPPGVNGTGGKH 120
Db 61 GPPKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPPGVNGTGGKH 120
Qy 121 GTPKKGPPKKGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Db 121 GTPKKGPPKKGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Qy 181 NASSGKFCVCGVPGIYFTYDTTLANKHLAIGLVNNGQYRITPDANTGNHVDVAGSTTLA 240
Db 181 NASSGKFCVCGVPGIYFTYDTTLANKHLAIGLVNNGQYRITPDANTGNHVDVAGSTTLA 240
Qy 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285
Db 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285

RESULT 4
US-10-621-787-2
; Sequence 2, Application US/10621787
; Publication No. US20040024187A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG
; FILE REFERENCE: 99-08D1
; CURRENT APPLICATION NUMBER: US/10/621,787
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: US 09/552,204
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: US 60/130,207
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-787-2

Query Match 100.0%; Score 1582; DB 15; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.4e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPVLLAALPCADPLGAFARDFRKS POLVCSLPGPQPPGPPGAPGPGMMGRM 60
Db 1 MIPVLLAALPCADPLGAFARDFRKS POLVCSLPGPQPPGPPGAPGPGMMGRM 60
Qy 61 GPPKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPPGVNGTGGKH 120
Db 61 GPPKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPPGVNGTGGKH 120
Qy 121 GTPKKGPPKKGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Db 121 GTPKKGPPKKGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Qy 181 NASSGKFCVCGVPGIYFTYDTTLANKHLAIGLVNNGQYRITPDANTGNHVDVAGSTTLA 240
Db 181 NASSGKFCVCGVPGIYFTYDTTLANKHLAIGLVNNGQYRITPDANTGNHVDVAGSTTLA 240
Qy 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285
Db 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285

RESULT 5
US-10-264-049-2619
; Sequence 2619, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
```

```
APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2619
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2619

Query Match 100.0%; Score 1582; DB 15; Length 314;
Best Local Similarity 100.0%; Pred. No. 1.6e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPVLLAALPCADPLGAFARDFRKS POLVCSLPGPQPPGPPGAPGPGMMGRM 60
Db 30 MIPVLLAALPCADPLGAFARDFRKS POLVCSLPGPQPPGPPGAPGPGMMGRM 89
Qy 61 GPPKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPPGVNGTGGKH 120
Db 90 GPPKQDQDGDHGRGSGSEGGPPGRTGNRKGPPKGAAGIAGAPGPPGVNGTGGKH 149
Qy 121 GTPKKGPPKKGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 180
Db 150 GTPKKGPPKKGKGGPGLPGPCSCSGGHTKSAFSAVATKSYRERLPIKFDKILNNEGHHY 209
Qy 181 NASSGKFCVCGVPGIYFTYDTTLANKHLAIGLVNNGQYRITPDANTGNHVDVAGSTTLA 240
Db 210 NASSGKFCVCGVPGIYFTYDTTLANKHLAIGLVNNGQYRITPDANTGNHVDVAGSTTLA 269
Qy 241 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 285
Db 270 LKQGDVWLQIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEY 314

RESULT 6
US-10-236-055A-6
; Sequence 6, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Matsson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Charles
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-055A-6

Query Match 100.0%; Score 1582; DB 14; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.7e-112;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPVLLAALPCADPLGAFARDFRKS POLVCSLPGPQPPGPPGAPGPGMMGRM 60
```

Db 46 MIPWVLLACALPCAADPLGAFARBRFRKSPOLVCSLPGPQBPBPBGAPGPGSGMMGRM 105  
Qy 61 GPRKQDGDHDDRGDSEBGPRTGNRGKPPKPKAGAIIGRAGPRGKGVNCTGKH 120  
Db 106 GPRKQDGDHDDRGDSEBGPRTGNRGKPPKPKAGAIIGRAGPRGKGVNCTGKH 165  
Qy 121 GTPKKGPKKKGKPGPLPGPCSCSGHTKSAFSAVATKSYPRERLPKFDKILMNEGGHY 180  
Db 166 GTPKKGPKKKGKPGPLPGPCSCSGHTKSAFSAVATKSYPRERLPKFDKILMNEGGHY 225  
Qy 181 NASSGKFCVCGVPPIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Db 226 NASSGKFCVCGVPPIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 285  
Qy 241 LKQDDEVWLQIFYSQNGLFYDPYWTDSLFTGFLIYADODDPNEV 285  
Db 286 LKQDDEVWLQIFYSQNGLFYDPYWTDSLFTGFLIYADODDPNEV 330

## RESULT 7

US-10-162-335-30  
; Sequence 30, Application US/10162335  
; Publication No. US20040009480A1  
GENERAL INFORMATION:  
; APPLICANT: Anderson, David W.  
; APPLICANT: Baumgartner, Jason C.  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Gangoli, Baha A.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Gorman, Linda  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Hjal, Tord  
; APPLICANT: Kehuda, Rameah  
; APPLICANT: Li, Li  
; APPLICANT: Macdougall, John R.  
; APPLICANT: Malyanekar, Uriel M.  
; APPLICANT: Miller, Isabelle  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Paturajan, Meera  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Stone, David J.  
; APPLICANT: Szytek, Kimberly A.  
; APPLICANT: Vernet, Corine A. M.  
; APPLICANT: Voss, Edward Z.  
; APPLICANT: Zethusen, Bryan D.  
TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method  
FILE REFERENCE: 21402-377 B  
CURRENT APPLICATION NUMBER: US/10/162,335  
PRIOR FILING DATE: 2002-10-01  
PRIOR APPLICATION NUMBER: 60/295,607  
PRIOR FILING DATE: 2001-06-04  
PRIOR APPLICATION NUMBER: 60/295,661  
PRIOR FILING DATE: 2001-06-04  
PRIOR APPLICATION NUMBER: 60/296,404  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: 60/296,418  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: 60/297,414  
PRIOR FILING DATE: 2001-06-11  
PRIOR APPLICATION NUMBER: 60/297,567  
PRIOR FILING DATE: 2001-06-12  
PRIOR APPLICATION NUMBER: 60/298,285  
PRIOR FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: 60/298,556  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: 60/299,949  
PRIOR FILING DATE: 2001-06-21  
PRIOR APPLICATION NUMBER: 60/300,883  
PRIOR FILING DATE: 2001-06-26

; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 201  
; SEQ ID NO 30  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-162-335-30

Query Match 100.0%; Score 1582; DB 15; Length 330;  
Best Local Similarity 100.0%; Pred. No. 1.7e-112;  
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPWVLLACALPCAADPLGAFARBRFRKSPOLVCSLPGPQBPBPBGAPGPGSGMMGRM 60  
Db 46 MIPWVLLACALPCAADPLGAFARBRFRKSPOLVCSLPGPQBPBPBGAPGPGSGMMGRM 105  
Qy 61 GPRKQDGDHDDRGDSEBGPRTGNRGKPPKPKAGAIIGRAGPRGKGVNCTGKH 120  
Db 106 GPRKQDGDHDDRGDSEBGPRTGNRGKPPKPKAGAIIGRAGPRGKGVNCTGKH 165  
Qy 121 GTPKKGPKKKGKPGPLPGPCSCSGHTKSAFSAVATKSYPRERLPKFDKILMNEGGHY 180  
Db 166 GTPKKGPKKKGKPGPLPGPCSCSGHTKSAFSAVATKSYPRERLPKFDKILMNEGGHY 225  
Qy 181 NASSGKFCVCGVPPIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 240  
Db 226 NASSGKFCVCGVPPIYFTYDITLANKHLAIGLVHNGQYRIRTPDANTGNHDVAGSTTLA 285  
Qy 241 LKQDDEVWLQIFYSQNGLFYDPYWTDSLFTGFLIYADODDPNEV 285  
Db 286 LKQDDEVWLQIFYSQNGLFYDPYWTDSLFTGFLIYADODDPNEV 330

## RESULT 8

US-10-220-120-354  
; Sequence 354, Application US/10220120  
; Publication No. US20040048253A1  
GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: PANZER, Scott R.  
; APPLICANT: SPIRO, Peter A.  
; APPLICANT: BANVILLE, Steven C.  
; APPLICANT: SHAH, Purvi  
; APPLICANT: CHALUP, Michael S.  
; APPLICANT: CHANG, Simon C.  
; APPLICANT: CHEN, Alice  
; APPLICANT: D'SA, Steven A.  
; APPLICANT: AMSEY, Stefan  
; APPLICANT: DAHL, Christopher R.  
; APPLICANT: DAM, Tam C.  
; APPLICANT: DANIELS, Susan E.  
; APPLICANT: DUPOUR, Gerard B.  
; APPLICANT: FLORES, Vincent  
; APPLICANT: FONG, Willy T.  
; APPLICANT: GREENAWALT, Lila B.  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: JONES, Antisa L.  
; APPLICANT: LIU, Tommy F.  
; APPLICANT: ROSEBERRY, Ann M.  
; APPLICANT: ROSEN, Bruce H.  
; APPLICANT: RUSSO, Frank D.  
; APPLICANT: STOCKDREHER, Theresa K.  
; APPLICANT: DAFO, Abel  
; APPLICANT: WRIGHT, Rachel J.  
; APPLICANT: YAP, Pierre B.  
; APPLICANT: YU, Jimmy Y.  
; APPLICANT: BRADLEY, Diana L.  
; APPLICANT: BRATCHER, Shawn R.  
; APPLICANT: CHEN, Wensheng  
; APPLICANT: COHEN, Howard J.  
; APPLICANT: HODGSON, David M.  
; APPLICANT: LINCOLN, Stephen E.  
; APPLICANT: JACKSON, Stuart

TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: PT-1113 PCT  
CURRENT APPLICATION NUMBER: US/10/220,120  
CURRENT FILING DATE: 2002-08-26  
PRIOR APPLICATION NUMBER: 60/184,771; 60/184,797; 60/184,698; 60/184,770; 60/184,774;  
60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;  
60/184,769; 60/184,768; 60/184,837; 60/184,841;  
60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;  
60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;  
60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;  
60/205,324; 60/205,286  
PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;  
2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;  
2000-05-17; 2000-05-17  
NUMBER OF SEQ ID NOS: 422  
SOFTWARE: PERL Program  
SEQ ID NO 354  
LENGTH: 331  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:414307.1.orf2:2000PEB01  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 191  
OTHER INFORMATION: unknown or other  
US-10-220-120-354

Query Match 99.7%; Score 1578; DB 15; Length 331;  
Best Local Similarity 99.6%; Pred. No. 3,3e-112;  
Matches 284; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIPVLLAALPCADPLTGAFAARDPRKSPOLVCSIPGQSPGPPGAGSPGGMGRM 60  
DB 47 MIPVLLAALPCADPLTGAFAARDPRKSPOLVCSIPGQSPGPPGAGSPGGMGRM 106  
QY 61 GPGKDGDDGHDGRSGSEGGPPGRTGNRKGPPKGAIGRAGPPGKVGTPGKH 120  
DB 107 GPGKDGDDGHDGRSGSEGGPPGRTGNRKGPPKGAIGRAGPPGKVGTPGKH 166  
QY 121 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNNEGHY 180  
DB 167 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNNEGHY 226  
QY 181 NASGKRVCGVPGIYFTYDITLANKHLAIGLVHNGQRIITPDANTGNHVAAGSTTLA 240  
DB 227 NASGKRVCGVPGIYFTYDITLANKHLAIGLVHNGQRIITPDANTGNHVAAGSTTLA 286  
QY 241 LKQGEVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADQDDPNEY 285  
DB 287 LKQGEVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADQDDPNEY 331

RESULT 9  
US-10-236-055A-8  
Sequence 8, Application US/10236055A  
Publication No. US20030134328A1  
GENERAL INFORMATION:  
APPLICANT: Basham, Beth E.  
APPLICANT: Forsythe, Ian  
APPLICANT: Gorman, Daniel M.  
APPLICANT: Mattson, Jeanne  
APPLICANT: Moshrefi, Christel  
TITLE OF INVENTION: MAMMALIAN GENES, RELATED REAGENTS  
FILE REFERENCE: DX01343K  
CURRENT APPLICATION NUMBER: US/10/236,055A  
CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/317,988  
PRIOR FILING DATE: 2001-09-06  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: Patent In version 3.1  
SEQ ID NO 8  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-236-055A-8

Query Match 94.3%; Score 1492; DB 14; Length 294;  
Best Local Similarity 94.4%; Pred. No. 1e-105;  
Matches 269; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 MIPVLLAALPCADPLTGAFAARDPRKSPOLVCSIPGQSPGPPGAGSPGGMGRM 60  
DB 10 MISVLLAALPCADPLTGAFAARDPRKSPOLVCSIPGQSPGPPGAGSGVGRM 69  
QY 61 GPGKDGDDGHDGRSGSEGGPPGRTGNRKGPPKGAIGRAGPPGKVGTPGKH 120  
DB 70 GPGKDGDDGHDGRSGSEGGPPGRTGNRKGPPKGAIGRAGPPGKVGTPGKH 129  
QY 121 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNNEGHY 180  
DB 130 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNNEGHY 189  
QY 181 NASGKRVCGVPGIYFTYDITLANKHLAIGLVHNGQRIITPDANTGNHVAAGSTTLA 240  
DB 190 NASGKRVCGVPGIYFTYDITLANKHLAIGLVHNGQRIITPDANTGNHVAAGSTTLA 249  
QY 241 LKQGEVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADQDDPNEY 285  
DB 250 LKQGEVWLQIFYSEONGLFYDPYWTDSLFTGFLIYADQDDPNEY 294

RESULT 10  
US-09-866-050A-294  
Sequence 294, Application US/09866050A  
Publication No. US20030040471A1  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Jorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Ornst, Rene  
APPLICANT: Murlson, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions isolated from skin cells  
FILE REFERENCE: 11000.1011c4V  
CURRENT APPLICATION NUMBER: US/09/866,050A  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 725  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 294  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Rat  
US-09-866-050A-294

Query Match 94.0%; Score 1487; DB 10; Length 294;  
Best Local Similarity 94.0%; Pred. No. 2,5e-105;  
Matches 268; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 MIPVLLAALPCADPLTGAFAARDPRKSPOLVCSIPGQSPGPPGAGSPGGMGRM 60  
DB 10 MISVLLAALPCADPLTGAFAARDPRKSPOLVCSIPGQSPGPPGAGSGVGRM 69  
QY 61 GPGKDGDDGHDGRSGSEGGPPGRTGNRKGPPKGAIGRAGPPGKVGTPGKH 120  
DB 70 GPGKDGDDGHDGRSGSEGGPPGRTGNRKGPPKGAIGRAGPPGKVGTPGKH 129  
QY 121 GTPGKGPCKGKGGPGLPGPCSCSGHTKSAFSAVATKSYRERLPPIKDKILNNEGHY 180

Db 130 GIPGKGGPKGKGGPGLPGPCSCSSRAKSAFSAVATKSYPRERLPIKFDKILMNEGHHY 189  
Qy 181 MASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGOYRIPTDPANTGNHDVAGSTTLA 240  
Db 190 MASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGOYRIPTDPANTGNHDVAGSTTLA 249  
Qy 241 LKQGDVWMLQIFYSBQNGLFYDPYMTDSLFTGFLIYADODDPNEV 285  
Db 250 LKQGDVWMLQIFYSBQNGLFYDPYMTDSLFTGFLIYADODDPNEV 294

RESULT 11  
US-09-866-050A-382

Sequence 382, Application US/09866050A  
Publication No. US2003004047A1  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions Isolated From Skin Cells  
TITLE OF INVENTION: and Methods for Their Use  
FILE REFERENCE: 11000.1011c4U  
CURRENT APPLICATION NUMBER: US/09/866,050A  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 725  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 382  
LENGTH: 285  
TYPE: PRT  
ORGANISM: Rat  
US-09-866-050A-382

Query Match 93.8%; Score 1484; DB 10; Length 285;  
Best Local Similarity 93.7%; Pred. No. 4,1e-105;  
Matches 267; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MIPWVLLACALPCAADPLGAFARDFRKSPOLVCSLPQGPBPBGAPGSGMMGRM 60  
Db 1 MISWMLACALPCAADPLGAFARDFRKSPOLVCSLPQGPBPBGAPGSGMMGRM 60  
Qy 61 GPFKQGDQDDHGRGDSGSEGGPRGTGNRGKPKKAGALGRAGRGPGVNGTGKH 120  
Db 61 GPFKQGDQDDHGRGDSGSEGGPRGTGNRGKPKKAGALGRAGRGPGVNGTGKH 120  
Qy 121 GTPGKGGPKGKGGPGLPGPCSCSSGHTKSAFSAVATKSYPRERLPIKFDKILMNEGHHY 180  
Db 121 GTPGKGGPKGKGGPGLPGPCSCSSGHTKSAFSAVATKSYPRERLPIKFDKILMNEGHHY 180  
Qy 181 MASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGOYRIPTDPANTGNHDVAGSTTLA 240  
Db 181 MASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGOYRIPTDPANTGNHDVAGSTTLA 240  
Qy 241 LKQGDVWMLQIFYSBQNGLFYDPYMTDSLFTGFLIYADODDPNEV 285  
Db 241 LKQGDVWMLQIFYSBQNGLFYDPYMTDSLFTGFLIYADODDPNEV 285

RESULT 12  
US-09-770-906-2

Sequence 2, Application US/09770906  
Patent No. US2001000905A1  
GENERAL INFORMATION:  
APPLICANT: Hensley, Preston  
APPLICANT: Hu, Erding  
APPLICANT: Smith, Randall, Forrest  
APPLICANT: Zhu, Yuan  
TITLE OF INVENTION: A Homolog of ACRP30 (30 KD  
TITLE OF INVENTION: ADIPOCTE COMPLEMENT-RELATED PROTEIN)  
FILE REFERENCE: GP-70435-C1  
CURRENT APPLICATION NUMBER: US/09/770,906

CURRENT FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/086,562  
PRIOR FILING DATE: 1998-05-21  
PRIOR APPLICATION NUMBER: 09/162,352  
PRIOR FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 217  
TYPE: PRT  
ORGANISM: HOMO SAPIENS  
US-09-770-906-2

Query Match 72.9%; Score 1154; DB 9; Length 217;  
Best Local Similarity 76.1%; Pred. No. 4,3e-80;  
Matches 217; Conservative 0; Mismatches 0; Indels 68; Gaps 1;

Qy 1 MIPWVLLACALPCAADPLGAFARDFRKSPOLVCSLPQGPBPBGAPGSGMMGRM 60  
Db 1 MISWMLACALPCAADPLGAFARDFRKSPOLVCSLPQGPBPBGAPGSGMMGRM 60  
Qy 61 GPFKQGDQDDHGRGDSGSEGGPRGTGNRGKPKKAGALGRAGRGPGVNGTGKH 120  
Db 61 GPFKQGDQDDHGRGDSGSEGGPRGTGNRGKPKKAGALGRAGRGPGVNGTGKH 87  
Qy 121 GTPGKGGPKGKGGPGLPGPCSCSSGHTKSAFSAVATKSYPRERLPIKFDKILMNEGHHY 180  
Db 88 -----VTKSYPRERLPIKFDKILMNEGHHY 112  
Qy 181 MASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGOYRIPTDPANTGNHDVAGSTTLA 240  
Db 113 MASSGKFCVCGVPGIYFTYDITLANKHLAIGLVHNGOYRIPTDPANTGNHDVAGSTTLA 172  
Qy 241 LKQGDVWMLQIFYSBQNGLFYDPYMTDSLFTGFLIYADODDPNEV 285  
Db 173 LKQGDVWMLQIFYSBQNGLFYDPYMTDSLFTGFLIYADODDPNEV 217

RESULT 13  
US-09-866-050A-138

Sequence 138, Application US/09866050A  
Publication No. US2003004047A1  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions Isolated From Skin Cells  
TITLE OF INVENTION: and Methods for Their Use  
FILE REFERENCE: 11000.1011c4U  
CURRENT APPLICATION NUMBER: US/09/866,050A  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 725  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 138  
LENGTH: 198  
TYPE: PRT  
ORGANISM: Rat  
US-09-866-050A-138

Query Match 62.0%; Score 981; DB 10; Length 198;  
Best Local Similarity 92.6%; Pred. No. 5,9e-67;  
Matches 175; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MIPWVLLACALPCAADPLGAFARDFRKSPOLVCSLPQGPBPBGAPGSGMMGRM 60  
Db 10 MISWMLACALPCAADPLGAFARDFRKSPOLVCSLPQGPBPBGAPGSGMMGRM 69  
Qy 61 GPFKQGDQDDHGRGDSGSEGGPRGTGNRGKPKKAGALGRAGRGPGVNGTGKH 120  
Db 70 GPFKQGDQDDHGRGDSGSEGGPRGTGNRGKPKKAGALGRAGRGPGVNGTGKH 129



QY 121 GTPGKGPKGKGGPGLPGPCSGSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGGHY 180  
 DB 130 GIPKGPKGKGGPGLPGPCSGSGHTKSAFSAVATKSYPRERLPIKFDKILNNEGGHY 189  
 QY 181 NASSGKFFVC 189  
 DB 190 NASSGKFFVC 198

## RESULT 14

US-10-234-000-15  
 ; Sequence 15, Application US/10234000  
 ; Publication No. US20030129698A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Piddington, Christopher S.  
 ; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN HOMOLOG ZACRP7  
 ; FILE REFERENCE: 99-31  
 ; CURRENT APPLICATION NUMBER: US/10/234,000  
 ; CURRENT FILING DATE: 2002-08-30  
 ; PRIOR APPLICATION NUMBER: US/09/577,298  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/136,289  
 ; PRIOR FILING DATE: 1999-05-27  
 ; PRIOR APPLICATION NUMBER: 60/145,589  
 ; PRIOR FILING DATE: 1999-07-22  
 ; PRIOR APPLICATION NUMBER: 60/158,448  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 15  
 ; LENGTH: 289  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-10-234-000-15

Query Match 56.8%; Score 898.5; DB 14; Length 289;

Best Local Similarity 62.7%; Pred. No. 1.7e-60; Matches 163; Conservative 32; Mismatches 62; Indels 3; Gaps 1;

QY 24 RRPFRKG---SPQLVCSLPPOGPPGPPGAPGSGMMGRMGFPKGQDGDHGDGDSGE 80  
 DB 19 RANQAKCESYSPRYICISIPGLPGPPGANGSGPHGRIGLPGDRDGRKSGKSGKT 78  
 QY 81 EGPPRTGNRGKPPKKAAGIAGAPRGPGVNGTPEKHGTPEKKGPPKGGEPGLPGP 140  
 DB 79 AGLKGTGPIGLAGEKGDQSGTGKGPILGPEGEKGVGAPGPPGKGRDQDGPGLPGV 138  
 QY 141 CSGSGGHTKSAFSAVATKSYPRERLPIKFDKILNNEGGHYASGKFFVCVPGIYFTYD 200  
 DB 139 CRGGSYLKSAFSAVATKSYPRERLPIKFDKILNNEGGHYASGKFFVCVPGIYFTYD 198  
 QY 201 ITLANKHLAIGLVNNGQYRIRTPDANTGNHVDVAGSTTLAKOGDEYWLQIFYSEONGLF 260  
 DB 199 ITLANKHLAIGLVNNGQYRIRTPDANTGNHVDVAGSTIVYILQPRDEVWLEIFPDQNGLF 258  
 QY 261 YDPYMTSLFTGFLIYADQD 280  
 DB 259 SDPGMADSLFSGFLIYVTD 278

## RESULT 15

US-09-866-050A-630  
 ; Sequence 630, Application US/09866050A  
 ; Publication No. US20030040471A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Watson, James D.  
 ; APPLICANT: Strachan, Lorna  
 ; APPLICANT: Sleeman, Matthew  
 ; APPLICANT: Onrust, Rene  
 ; APPLICANT: Muriel, James G.  
 ; APPLICANT: Kumble, Kriehand D.

; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
 ; FILE REFERENCE: 11000.1011c4U  
 ; CURRENT APPLICATION NUMBER: US/09/866,050A  
 ; NUMBER OF SEQ ID NOS: 725  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 630  
 ; LENGTH: 289  
 ; TYPE: PRT  
 ; ORGANISM: Rat  
 ; US-09-866-050A-630

Query Match 56.8%; Score 898; DB 10; Length 289;

Best Local Similarity 64.0%; Pred. No. 1.9e-60; Matches 160; Conservative 31; Mismatches 59; Indels 0; Gaps 0;

QY 31 SPQLVCSLPPOGPPGPPGAPGSGMMGRMGFPKGQDGDHGDGDSGEGPPRTGNR 90  
 DB 29 SPRYICISIPGLPGPPGANGSGPHGRIGLPGDRDGRKSGKGTAGLKTGPL 88  
 QY 91 GKEPPKKAAGIAGAPRGPGVNGTPEKHGTPEKKGPPKGGEPGLPGPCSGSGHTKS 150  
 DB 89 GLAGEKGDQSGTGKGPILGPEGEKGVGAPGPPGKGRDQDGPGLPGVCRGSIYLS 148  
 QY 151 AFSVAATKSYPRERLPIKFDKILNNEGGHYASGKFFVCVPGIYFTYDITLANKHLAI 210  
 DB 149 AFSVATKSYPRERLPIKFDKILNNEGGHYASGKFFVCVPGIYFTYDITLANKHLAI 208  
 QY 211 GLVNGQYRIRTPDANTGNHVDVAGSTTLAKOGDEYWLQIFYSEONGLFYDPYMTSLF 270  
 DB 209 GLVNGQYRIRTPDANTGNHVDVAGSTIVYILQPRDEVWLEIFPDQNGLSDPGMADSLF 268  
 QY 271 TGFLLIYADQD 280  
 DB 269 SGFLIYVTD 278

Search completed: December 25, 2004, 09:22:17

Job time : 505.456 secs

This Page Blank (USD)  
Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 25, 2004, 06:40:18 ; Search time 33.6487 Seconds

(without alignments)  
330.491 Million cell updates/sec

Title: US-10-621-787-5

Perfect score: 57  
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXGXXFX 31Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 100 summaries

## Database :

1: A\_Geneseq\_23Sep04:\*  
2: geneseqp1980s:\*  
3: geneseqp1990s:\*  
4: geneseqp2000s:\*  
5: geneseqp2001s:\*  
6: geneseqp2002s:\*  
7: geneseqp2003as:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	22	38.6	119 8 ADK66184	Adk66184 Mouse phx
2	22	38.6	136 8 ADK66189	Adk66189 Human phx
3	22	38.6	163 2 AAW70458	AAW70458 Human T1-
4	22	38.6	188 4 AAM93624	Aam93624 Human pol
5	22	38.6	188 8 ADL1427	Adl1427 Human pro
6	22	38.6	215 2 AAW70459	AAW70459 Human T1-
7	22	38.6	215 5 ABP65162	ABP65162 Hypoxia-r
8	22	38.6	219 4 ABB11780	Abb11780 Human T1
9	22	38.6	224 4 AAM93844	Aam93844 Human pol
10	22	38.6	224 4 AAB92820	Aab92820 Human pro
11	22	38.6	224 4 ADL31890	Adl31890 Human pro
12	22	38.6	231 6 ABR41674	Ab41674 Human DIT
13	22	38.6	266 4 AAG82486	Aag82486 S. epider
14	22	38.6	268 7 ABO69527	ABO69527 Pseudomon
15	22	38.6	292 5 ABP39297	ABP39297 Staphyloc
16	22	38.6	296 6 ABU39079	Abu39079 Protein e
17	22	38.6	302 6 ABU49473	Abu49473 Protein e
18	22	38.6	305 6 ADA34782	Ada34782 Actinobac
19	22	38.6	404 6 ADA36895	Ada36895 Actinobac
20	22	38.6	485 6 ABM67542	Abm67542 Photocarb
21	22	38.6	530 8 ADP04497	Adp04497 Sea equit
22	22	38.6	540 4 AAU57054	Aau57054 Propionib
23	22	38.6	540 6 ABM53573	Abm53573 Proteinib
24	22	38.6	3705 6 ABU50020	Abu50020 Protein e
25	21	36.8	20 4 AAB73384	Aab73384 T. harzia

26	21	36.8	53 8 ABO55435	ABO55435 Human gen
27	21	36.8	111 2 AAW89961	AAW89961 Antigen f
28	21	36.8	128 5 ABG76563	ABG76563 HCV El an
29	21	36.8	151 2 AAY37071	Aay37071 Protein w
30	21	36.8	154 6 ABP75824	ABP75824 Human sec
31	21	36.8	156 2 AAY35280	Aay35280 Chlamydia
32	21	36.8	156 6 ABU26784	Abu26784 Protein e
33	21	36.8	161 6 ABM69553	Abm69553 Photocarb
34	21	36.8	169 5 ABP27601	ABP27601 Streptoco
35	21	36.8	183 5 ABP69421	ABP69421 Human pol
36	21	36.8	201 4 AAB74750	Aab74750 Human sec
37	21	36.8	201 4 AAB74759	Aab74759 Human sec
38	21	36.8	201 4 AAB74760	Aab74760 Human sec
39	21	36.8	201 5 ABG65327	ABG65327 Human alb
40	21	36.8	201 5 ABG65326	ABG65326 Human alb
41	21	36.8	201 5 ABG65324	ABG65324 Human alb
42	21	36.8	201 7 ADD44752	Add44752 Human pro
43	21	36.8	201 7 ADD44746	Add44746 Rat Prote
44	21	36.8	201 7 ADD44750	Add44750 Rat Prote
45	21	36.8	201 7 ADD44748	Add44748 Human pro
46	21	36.8	201 7 ADJ68911	Adj68911 Human hea
47	21	36.8	201 7 ADP65317	ADP65317 Human coa
48	21	36.8	201 8 ADL78593	Adl78593 Albumin f
49	21	36.8	201 8 ADL78591	Adl78591 Albumin f
50	21	36.8	201 8 ADL78594	Adl78594 Albumin f
51	21	36.8	201 8 ADN04050	Adn04050 Antipepti
52	21	36.8	208 5 ABP69422	ABP69422 Human pol
53	21	36.8	210 2 AAW20426	AAW20426 H. pylori
54	21	36.8	215 2 AAW20968	AAW20968 H. pylori
55	21	36.8	238 6 ABU30638	Abu30638 Protein e
56	21	36.8	252 7 ABO82067	ABO82067 Pseudomon
57	21	36.8	254 4 AAU31151	AAU31151 Novel hum
58	21	36.8	254 7 ADB10062	ADB10062 Novel pro
59	21	36.8	271 3 AAG51318	Aag51318 Arabidops
60	21	36.8	271 3 AAG07614	Aag07614 Arabidops
61	21	36.8	277 3 AAG51317	Aag51317 Arabidops
62	21	36.8	277 3 AAG07613	Aag07613 Arabidops
63	21	36.8	277 7 ADD30338	Add30338 Plant yie
64	21	36.8	277 8 ADI44261	ADI44261 Plant tra
65	21	36.8	280 5 ABP29735	ABP29735 Streptoco
66	21	36.8	281 4 ABG23769	ABG23769 Novel hum
67	21	36.8	295 3 AAG51316	Aag51316 Arabidops
68	21	36.8	296 3 AAG07612	Aag07612 Arabidops
69	21	36.8	300 3 AAY91317	Aay91317 Group B S
70	21	36.8	300 5 ABP26518	ABP26518 Streptoco
71	21	36.8	307 3 AAB18332	Aab18332 Plasmodin
72	21	36.8	336 5 ABP62884	ABP62884 Human pol
73	21	36.8	348 7 ADM25554	Adm25554 Hyperther
74	21	36.8	382 4 ABB60990	Abb60990 Drosophi
75	21	36.8	408 4 AAB79473	Aab79473 Corynebac
76	21	36.8	441 4 AAG92180	Aag92180 C glutami
77	21	36.8	450 7 ABO63535	ABO63535 Klebsiell
78	21	36.8	473 5 AAU72909	Aau72909 Human met
79	21	36.8	473 5 ABB98126	Abb98126 Human PPM
80	21	36.8	508 3 AAY77945	Aay77945 A. thalia
81	21	36.8	680 3 AAG45662	Aag45662 Arabidops
82	21	36.8	729 3 AAG45661	Aag45661 Arabidops
83	21	36.8	783 3 AAG45660	Aag45660 Arabidops
84	21	36.8	803 4 ABB64390	Abb64390 Drosophi
85	21	36.8	1032 4 ABB62164	Abb62164 Drosophi
86	21	36.8	1273 7 ADH87086	Adh87086 Enterococ
87	21	36.8	1416 2 AAY00211	Aay00211 Enterococ
88	21	36.8	1416 5 ABP43430	ABP43430 E faecali
89	21	36.8	1416 6 ABU88458	Abu88458 E. faecali
90	21	36.8	1416 6 ABU13709	Abu13709 Enterococ
91	21	36.8	1448 6 AAY00210	Aay00210 Enterococ
92	21	36.8	1448 5 ABP43429	ABP43429 E faecali
93	21	36.8	1448 6 ABU88457	Abu88457 E. faecali
94	21	36.8	1448 6 ABU13708	Abu13708 Enterococ
95	20	35.1	53 4 AAU21262	Aau21262 Human nov
96	20	35.1	55 8 ABO54754	ABO54754 Human gen
97	20	35.1	79 3 AAG22312	Aag22312 Zea may
98	20	35.1	94 3 AAG22311	Aag22311 Zea may

99 20 35.1 118 4 AAm94043  
100 20 35.1 122 8 ADn47176

AAm94043 Human rep  
ADn47176 Thermococ

## ALIGNMENTS

## RESULT 1

ADK66184  
ID ADK66184 standard; protein; 119 AA.

AC ADK66184;

DT 06-MAY-2004 (first entry)

DE Mouse p1x-6 protein.

XX Novel gene; pain; neuropathic pain; shingles pain;

KM post-herpetic neuralgia; gene therapy; mouse.

XX Mus musculus.

OS Mus musculus.

XX Key Location/Qualifiers

FT Misc-difference 110 /note= "Encoded by TGC"

FT Misc-difference 117 /note= "Encoded by GTG"

XX US2004019006-A1.

XX 29-JAN-2004.

XX 09-MAY-2003; 2003US-00434156.

XX 10-MAY-2002; 2002US-0378955P.

XX (HAYA/) HAYASHIZAKI Y.

XX (KAMI/) KAMIYA M.

XX (SUZU/) SUZUKI T.

XX (HIRO/) HIROKAWA H.

XX (NAKA/) NAKAO K.

XX (TANA/) TANAKA T.

XX Hayashizaki Y, Kamiya M, Suzuki T, Hirokawa H, Nakao K, Tanaka T;

XX WPI; 2004-122083/12.

XX N-PSDB; ADK66173.

XX New isolated polynucleotide, useful for useful for treating or preventing

XX pain, e.g. neuropathic pain, preferably shingles pain or post-herpetic

XX neuralgia.

XX Example 9; SEQ ID NO 17; 53pp; English:

XX The present invention provides novel genes relating to pain. The

XX invention is useful for treating and preventing pain such as neuropathic

XX CC pain, shingles pain and post-herpetic neuralgia. The invention is also

XX useful in gene therapy. The present sequence is mouse p1x-6 protein.

XX SQ Sequence 119 AA;

XX Query Match 38.6%; Score 22; DB 8; Length 119;

XX Best Local Similarity 36.4%; Pred. No. 2.8e+02;

XX Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 10 FTASRNGTYKF 20

RESULT 2

ADK66189

ID ADK66189 standard; protein; 136 AA.

XX ADK66189;

AC ADK66189;

DT 06-MAY-2004 (first entry)

DE Human p1x-6 protein.

XX Novel gene; pain; neuropathic pain; shingles pain;

KM post-herpetic neuralgia; gene therapy; human.

XX Homo sapiens.

XX US2004019006-A1.

XX 29-JAN-2004.

XX 09-MAY-2003; 2003US-00434156.

XX 10-MAY-2002; 2002US-0378955P.

XX (HAYA/) HAYASHIZAKI Y.

XX (KAMI/) KAMIYA M.

XX (SUZU/) SUZUKI T.

XX (HIRO/) HIROKAWA H.

XX (NAKA/) NAKAO K.

XX (TANA/) TANAKA T.

XX Hayashizaki Y, Kamiya M, Suzuki T, Hirokawa H, Nakao K, Tanaka T;

XX WPI; 2004-122083/12.

XX N-PSDB; ADK66178.

XX New isolated polynucleotide, useful for useful for treating or preventing

XX pain, e.g. neuropathic pain, preferably shingles pain or post-herpetic

XX neuralgia.

XX Claim 10; SEQ ID NO 22; 53pp; English.

XX The present invention provides novel genes relating to pain. The

XX invention is useful for treating and preventing pain such as neuropathic

XX CC pain, shingles pain and post-herpetic neuralgia. The invention is also

XX useful in gene therapy. The present sequence is human p1x-6 protein.

XX SQ Sequence 136 AA;

XX Query Match 38.6%; Score 22; DB 8; Length 136;

XX Best Local Similarity 36.4%; Pred. No. 3.2e+02;

XX Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 10 FTASRNGTYKF 20

RESULT 3

AAW70458

ID AAW70458 standard; protein; 163 AA.

XX AAW70458;

XX 10-DEC-1998 (first entry)

XX Human T1-receptor ligand III splice variant 1.

XX Human T1-receptor ligand III; T1-R ligand III; atherosclerosis;

XX autoimmune disease; inflammation; metabolic dysfunction;

XX immune-regulated disorder.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1.24

FT /note= "Signal peptide"

FT Protein 25.163  
 FT /note="Tl-R ligand III splice variant 1"  
 XX  
 XX MO9838311-A1.  
 XX  
 XX 03-SEP-1998.  
 XX  
 XX 26-FEB-1998; 98WO-US003483.  
 XX  
 XX 28-FEB-1997; 97US-0039483P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX N1 J, Gentz RL, Ruben SM;  
 XX  
 XX WPI; 1998-495398/42.  
 XX  
 XX N-PSDB; AAV33461.  
 XX  
 XX New human Tl-receptor ligand III splice variant (s) - useful for, e.g.  
 PT detecting and treating immune system related disorders such as cancer and  
 PT inflammation.  
 PS Claim 1; Fig 1; 115pp; English.  
 CC The invention provides novel human Tl-receptor ligand III (Tl-R ligand  
 CC III) splice variants and antibodies raised against these proteins. The  
 CC present sequence represents the Tl-R ligand III splice variant 1 protein  
 CC sequence. The Tl-R ligand III splice variants are claimed useful for  
 CC screening agonists and antagonists. They are also claimed useful for  
 CC treating disorders such as atherosclerosis, autoimmune disease,  
 CC inflammation, metabolic dysfunction and immune-regulated disorders  
 XX  
 XX Sequence 163 AA;

Query Match 38.6%; Score 22; DB 2; Length 163;  
 Best Local Similarity 36.4%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 19 FXXXXXGXKXKF 29  
 DB 89 FTASKNGTYKRF 99

RESULT 4  
 AAM93624  
 ID AAM93624 standard; protein; 188 AA.  
 XX  
 XX AAM93624;  
 AC  
 XX  
 XX 06-NOV-2001 (first entry)  
 DT  
 XX  
 XX Human polypeptide, SEQ ID NO: 3460.  
 DE  
 XX  
 XX Human, full length cDNA; cDNA synthesis; oligo-capping.  
 KM  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX EP130094-A2.  
 PN  
 XX  
 XX 05-SEP-2001.  
 PD  
 XX  
 XX 07-JUL-2000; 2000EP-00114089.  
 PF  
 XX  
 XX 08-JUL-1999; 99JP-00194486.  
 PR  
 XX  
 XX 11-JAN-2000; 2000JP-00118774.  
 PR  
 XX  
 XX 02-MAY-2000; 2000JP-00183765.  
 XX  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA  
 XX  
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX  
 XX WPI; 2001-524255/58.  
 DR

DR N-PSDB; AAK94557.  
 XX  
 XX 830 Primers useful for synthesizing full length cDNA clones and their use  
 PT in genetic manipulation.  
 FT  
 XX  
 XX Claim 8; SEQ ID NO 3460; 1380pp + Sequence Listing; English.  
 PS  
 XX  
 XX The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been isolated  
 CC and nucleotide sequences of 5' - and 3' - ends of the cDNA molecules have  
 CC been determined. Primers for synthesizing the full length cDNA are useful  
 CC for clarifying the function of the protein encoded by the cDNA. The full  
 CC length clones were obtained by construction of full length enriched cDNA  
 CC libraries that were synthesized by the oligo-capping method. The primers  
 CC enable the production of the full length cDNA easily without any special  
 CC methods. The present sequence is a polypeptide encoded by the full length  
 CC human cDNA of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in CD-ROM  
 CC format directly from EPO  
 XX  
 XX Sequence 188 AA;

Query Match 38.6%; Score 22; DB 4; Length 188;  
 Best Local Similarity 36.4%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 19 FXXXXXGXKXKF 29  
 DB 98 FTASKNGTYKRF 108

RESULT 5  
 ADL31427  
 ID ADL31427 standard; protein; 188 AA.  
 XX  
 XX ADL31427;  
 AC  
 XX  
 XX 20-MAY-2004 (first entry)  
 DT  
 XX  
 XX Human protein encoded by a full length cDNA clone SegID 3460.  
 DE  
 XX  
 XX human; medicine; signal transduction; glycoprotein; transcription;  
 KW oligo-capping method.  
 KW  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX EP1396543-A2.  
 PN  
 XX  
 XX 10-MAR-2004.  
 PD  
 XX  
 XX 07-JUL-2000; 2003EP-00025638.  
 PF  
 XX  
 XX 08-JUL-1999; 99JP-00194486.  
 PR  
 XX  
 XX 11-JAN-2000; 2000JP-00118774.  
 PR  
 XX  
 XX 02-MAY-2000; 2000JP-00183865.  
 PR  
 XX  
 XX 07-JUL-2000; 2000EP-00114089.  
 XX  
 XX  
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
 PA  
 XX  
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX  
 XX WPI; 2004-204755/20.  
 DR  
 XX  
 XX N-PSDB; ADL31426.  
 DR  
 XX  
 XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
 PT length human cDNAs.  
 PT  
 XX  
 XX Example 1; SEQ ID NO 3460; 1340pp; English.  
 PS  
 XX  
 XX This invention relates to a novel primers useful for synthesizing full  
 CC length cDNA molecules that encode human proteins. Specifically, it refers  
 CC to secretory or membrane proteins that are potential therapeutic agents/

CC target molecules in the field of medicine, and in particular genes  
 CC encoding proteins that are associated with signal transduction.  
 CC glycoproteins and transcription. The present invention describes a method  
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
 CC ends using the oligo-capping method. This polypeptide sequence is a full  
 CC length human protein of the invention.

XX Sequence 188 AA;

Query Match 38.6%; Score 22; DB 8; Length 188;

Best Local Similarity 36.4%; Pred. No. 4.4e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYX 29

Db 98 FTASKNGTYKF 108

#### RESULT 6

AAW70459 ID AAW70459 standard; protein; 215 AA.

AC AAW70459;

DT 10-DEC-1998 (first entry)

XX Human T1-receptor ligand III splice variant 2.

XX Human T1-receptor ligand III; T1-R ligand III; atherosclerosis;

KW autoimmune disease; inflammation; metabolic dysfunction;

KW immune-regulated disorder.

XX Homo sapiens.

OS Key Location/Qualifiers

PH Peptide 1..24

PT /note= "Signal peptide"

PT Protein 25..215

PT /note= "T1-R ligand III splice variant 2"

XX WO9838311-A1.

PN 03-SEP-1998.

PD 26-FEB-1998; 98WO-US003483.

PR 28-FEB-1997; 97US-0039483P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA N1 J, Gentz RL, Ruben SM;

PI WPI; 1998-495398/42.

DR N-PSDB; AAV33462.

XX New human T1-receptor ligand III splice variant(s) - useful for, e.g.

PT detecting and treating immune system related disorders such as cancer and

PT inflammation.

XX Claim 1; Fig 2; 115pp; English.

XX The invention provides novel human T1-receptor ligand III (T1-R ligand

CC III) splice variants and antibodies raised against these proteins. The

CC present sequence represents the T1-R ligand III splice variant 2 protein

CC sequence. The T1-R ligand III splice variants are claimed useful for

CC screening agonists and antagonists. They are also claimed useful for

CC treating disorders such as atherosclerosis, autoimmune disease,

CC inflammation, metabolic dysfunction and immune-regulated disorders

XX Sequence 215 AA;

Query Match 38.6%; Score 22; DB 2; Length 215;

Best Local Similarity 36.4%; Pred. No. 5e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYX 29

Db 89 FTASKNGTYKF 99

#### RESULT 7

ABP65162 ID ABP65162 standard; protein; 215 AA.

AC ABP65162;

DT 12-NOV-2002 (first entry)

XX Hypoxia-regulated protein #36.

XX Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;

KW antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;

KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;

KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;

KW preclapmia; atherosclerosis; inflammatory condition; wound healing;

KW inflammation; erythropoiesis; hair loss; human.

XX Homo sapiens.

OS WO200246465-A2.

PN 13-JUN-2002.

PD 10-DEC-2001; 2001WO-GB005458.

PR 08-DEC-2000; 2000GB-00030076.

PR 08-FEB-2001; 2001GB-00003156.

PR 25-OCT-2001; 2001GB-00025666.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

PA White J, Mundy CR, Ward NR, Krige D, Kingman SM, Harris RA;

PI Rayner MN;

PI WPI; 2002-627238/67.

DR 2002-627238/67.

PD Identifying a gene involved in disease for treating hypoxia-regulated

PT conditions, comprises comparing the transcriptome/proteome of two cell

PT types under different conditions and identifying a differentially

PT regulated gene.

XX Claim 35; Page 355; 538pp; English.

XX The present invention relates to methods for identifying genes and

CC proteins that are implicated in a specific disease or physiological

CC condition. The method comprises comparing the transcriptome/proteome of a

CC specialised cell type implicated in a disease or condition with that of a

CC second specialised cell type, under two experimental conditions, and

CC identifying a gene that is differentially regulated in the two

CC specialised cell types under experimental conditions. ABV7873-ABV78116

CC and ABP65061-ABP65257 were identified using the methods of the invention.

CC The coding sequences and proteins are useful for treating a disease in a

CC patient, for manufacture of a medicament for treating hypoxia-regulated

CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,

CC biological response to hypoxia conditions, or hypoxic-associated

CC pathology in a patient. The coding sequences and proteins are also useful

CC for monitoring the therapeutic treatment of a disease or physiological

CC condition, such as cancer, ischaemic conditions, reperfusion injury,

CC retinopathy, neonatal stress, preclapmia, atherosclerosis, inflammatory

CC conditions, wound healing, inflammation, erythropoiesis or hair loss

XX Sequence 215 AA;

Query Match 38.6%; Score 22; DB 5; Length 215;

Best Local Similarity 36.4%; Pred. No. 5e+02; Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;



QY 19 FXXXXXGXYF 29  
DB 90 FTASKNGTYKF 100

RESULT 8  
ABBI1780  
ID ABBI1780 standard; peptide: 219 AA.  
XX  
AC ABBI1780;  
XX  
DT 11-JAN-2002 (first entry)  
XX  
DE Human T1 receptor ligand III homologue, SEQ ID NO:2150.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; antiinflammatory;  
KW antiaesthetic; antiarthritic; haemostatic; antiarteriosclerotic;  
KW cytoskeletal; osteopathic; vasotropic; cardiac; virucide; antibacterial;  
KW antifungal; vulnery; antitumor.

OS Homo sapiens.  
XX  
PN W0200157188-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US003800.  
XX  
PR 03-FEB-2000; 2000US-00496914.  
XX  
PR 27-APR-2000; 2000US-00560875.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR N-PSDB; ABA09024.  
XX  
PT Human proteins and DNA encoding sequences useful for preventing, treating  
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis  
PT and cancer.  
XX  
PS Claim 20; Page 245; 1963pp; English.

CC Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibitor-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical

CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness.  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a novel human  
CC polypeptide of the invention

QY 19 FXXXXXGXYF 29  
DB 93 FTASKNGTYKF 103

Query Match 38.6%; Score 22; DB 4; Length 219;  
Best Local Similarity 36.4%; Pred. No. 5.1e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 9  
AAM93844  
ID AAM93844 standard; protein: 224 AA.  
XX  
AC AAM93844;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide, SEQ ID NO: 3923.  
XX  
KW Human; full length cDNA; cDNA synthesis; oligo-capping.  
XX  
OS Homo sapiens.  
XX  
PN EP1130094-A2.  
XX  
PD 05-SEP-2001.  
XX  
PF 07-JUL-2000; 2000EP-00114089.  
XX  
PR 08-JUL-1999; 99JP-00194486.  
XX  
PR 11-JAN-2000; 2000JP-00118774.  
XX  
PR 02-MAY-2000; 2000JP-00183765.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX  
DR N-PSDB; AAK94800.  
XX  
PT 830 Primers useful for synthesizing full length cDNA clones and their use  
PT in genetic manipulation.  
XX  
PS Claim 8; SEQ ID NO 3923; 1380pp + Sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been isolated  
CC and nucleotide sequences of 5' - and 3' - ends of the cDNA molecules have  
CC been determined. Primers for synthesizing the full length cDNA are useful  
CC for clarifying the function of the protein encoded by the cDNA. The full  
CC length clones were obtained by construction of full length enriched cDNA

CC libraries that were synthesised by the oligo-capping method. The primers  
 CC enable the production of the full length cDNA easily without any special  
 CC methods. The present sequence is a polypeptide encoded by a full length  
 CC human cDNA of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in CD-ROM  
 CC format directly from EPO

XX Sequence 224 AA;

Query Match 38.6%; Score 22; DB 4; Length 224;

Best Local Similarity 36.4%; Pred. No. 5.3e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYP 29

DB 98 FTASKNGTYKF 108

RESULT 10

AA92820 standard; protein; 224 AA.

AC AAB92820;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:11348.

KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-00116126.

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-

PT length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length

PT cDNAs.

PS Claim 8; SEQ ID NO 11348; 2537bp + Sequence listing; English.

XX The present invention describes primer sets for synthesising 5602 full-

XX length cDNAs defined in the specification. Where a primer set comprises:

XX (a) an oligo-dT primer and an oligonucleotide complementary to the

XX complementary strand of a polynucleotide which comprises one of the 5602

XX nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in the

XX specification. The primer sets can be used in antisense therapy and in

XX gene therapy. The primers are useful for synthesising polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893

CC represent human amino acid sequences; and AAH13629 to AAH13632 represent

CC oligonucleotides, all of which are used in the exemplification of the

CC present invention

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Sequence 224 AA;

Query Match 38.6%; Score 22; DB 4; Length 224;

Best Local Similarity 36.4%; Pred. No. 5.3e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYP 29

DB 98 FTASKNGTYKF 108

RESULT 11

ADL1890 standard; protein; 224 AA.

AC ADL1890;

DT 20-MAY-2004 (first entry)

DE Human protein encoded by a full length cDNA clone SegID 3923.

KM human; medicine; signal transduction; glycoprotein; transcription;

OS oligo-capping method.

PN EP1396543-A2.

PD 10-MAR-2004.

PF 07-JUL-2000; 2003EP-00025638.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183865.

PR 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2004-204755/20.

XX N-PSDB; ADL1889.

XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full

XX length human cDNAs.

PS Example 1; SEQ ID NO 3923; 1340bp; English.

XX This invention relates to a novel primers useful for synthesising full

XX length cDNA molecules that encode human proteins. Specifically, it refers

XX to secretory or membrane proteins that are potential therapeutic agents/

XX target molecules in the field of medicine, and in particular genes

XX encoding proteins that are associated with signal transduction.

XX glycoproteins and transcription. The present invention describes a method

XX for efficiently cloning a full length human cDNA from both the 5' and 3'

XX ends using the oligo-capping method. This polypeptide sequence is a full

XX length human protein of the invention.

XX Sequence 224 AA;

Query Match 38.6%; Score 22; DB 8; Length 224;

Best Local Similarity 36.4%; Pred. No. 5.3e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYP 29

DB 98 FTASKNGTYKF 108

QY 19 FXXXXGXXYP 29  
DB 98 FTASKNGTYKF 108

RESULT 12  
ID ABR41674 standard; protein; 231 AA.  
XX ABR41674;  
XX 02-JUN-2003 (first entry)  
DE Human DTHP cell membrane protein.  
XX Human; dthp; diagnostic and therapeutic polynucleotide; diagnosis;  
KM cancer; cell proliferative disorder; autoimmune disorder;  
KM inflammatory disorder; infection; hormonal disorder; metabolic disorder;  
KM neurological disorder; gastrointestinal disorder; transport disorder;  
KM connective tissue disorder; drug screening; proteome analysis;  
KM gene therapy; antisense therapy; genotyping; transgenic animal; knock in;  
KM disease model; toxicological testing; transcript imaging;  
KM cell membrane protein.

OS Homo sapiens.  
PN WO200297031-A2.  
PD 05-DEC-2002.  
PF 27-MAR-2002; 2002WO-US010056.  
XX 28-MAR-2001; 2001US-0279619P.  
PR 29-MAR-2001; 2001US-0280067P.  
PR 29-MAR-2001; 2001US-0280068P.  
PR 16-MAY-2001; 2001US-0291280P.  
PR 17-MAY-2001; 2001US-0291829P.  
PR 17-MAY-2001; 2001US-0291849P.  
PR 19-JUN-2001; 2001US-0299428P.  
PR 20-JUN-2001; 2001US-0299776P.  
PR 20-JUN-2001; 2001US-0300001P.  
XX (INCY-) INCTRE GENOMICS INC.  
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshery SR;  
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleeefeld Y, Gerstein EH;  
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B,  
PI Flores V, Marwaha R, Lo A, Lan RY, Uraahka ME;  
XX WPI; 2003-129518/12.  
DR N-PSDB; ACC46611.  
XX Novel human diagnostic and therapeutic polypeptide useful for identifying  
PT test compound which specifically binds to a polypeptide encoded by human  
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.  
XX Claim 27; SEQ ID NO 1209; 591pp; English.

XX The invention relates to novel human diagnostic and therapeutic  
CC polynucleotides designated dthp (ACC46080-ACC46749) and to their encoded  
CC proteins (DTHP; ABR41136-ABR41812). The invention also relates to  
CC polynucleotide sequences at least 90% identical to the dthp cDNA  
CC sequences of the invention; recombinant vectors, host cells and  
CC transgenic organisms comprising a dthp nucleic acid sequence; the  
CC recombinant production of DTHP proteins; antibodies specific for DTHP  
CC proteins; microarrays comprising dthp nucleic acid sequences; methods of  
CC detecting dthp nucleotide and protein sequences; methods of screening  
CC for compounds which specifically bind a DTHP protein; and methods of  
CC assessing the toxicity of test compounds using a dthp hybridisation  
CC probe. Dthp nucleic acid sequences and DTHP proteins may be used in the  
CC diagnosis of a wide variety of conditions including cancer and other cell

CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,  
CC viral, fungal or parasitic infections; hormonal disorders; metabolic  
CC disorders; neurological disorders; gastrointestinal disorders; transport  
CC disorders; and connective tissue disorders. They may also be used to  
CC screen for modulators of protein activity or gene expression. DTHP  
CC proteins can additionally be used in analysis of the proteome of a tissue  
CC or cell type and to induce antibodies. The dthp nucleic acids are  
CC additionally useful in somatic or germline gene therapy of the disorders  
CC mentioned above, as a source of antisense sequences, as a source of  
CC probes and primers, in genotyping and identification of individuals, in  
CC the generation of transgenic animal models of human disease or knock in  
CC humanised animals, in toxicological testing, and in transcript imaging.  
CC The present sequence represents a DTHP protein which is a cell membrane  
CC protein. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
XX

Seq Sequence 231 AA;  
QY 19 FXXXXGXXYP 29  
DB 105 FTASKNGTYKF 115

Query Match 38.6%; Score 22; DB 6; Length 231;  
Best Local Similarity 36.4%; Pred. No. 5.4e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 13  
ID AAG82486 standard; protein; 266 AA.  
XX AAG82486;  
XX 03-SEP-2001 (first entry)  
DE S. epidermidis open reading frame protein sequence SEQ ID NO:2066.  
XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
KM endocarditis.  
XX Staphylococcus epidermidis.  
OS  
PN WO200134809-A2.  
PD 17-MAY-2001.  
XX 09-NOV-2000; 2000WO-US030782.  
PR 09-NOV-1999; 99US-0164258P.  
XX (GLAX) GLAXO GROUP LTD.  
XX Kimmery WJ;  
XX WPI; 2001-316495/33.  
DR N-PSDB; AAG53336.  
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
PT useful for vaccinating against infections, e.g. endocarditis.  
XX Claim 18; Page 560; 2186pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
CC (II), given in AAG83454 to AAG83120, from Staphylococcus epidermidis. (I)  
CC and (II) can have antibacterial activity and therefore can be used in  
CC vaccination. The nucleic acids (I) may be used to produce the S.  
CC epidermidis polypeptides (II) via the production of vectors containing  
CC them which are used to produce host cells which express the  
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
CC used to vaccinate subjects and to raise antibodies against the bacteria.  
CC The polypeptides may also be used to assay for other inhibitors of their  
CC activity and therefore identify compounds that may be used for the

CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to  
CC AAH55090 represent specifically claimed *S. epidermidis* genomic DNA  
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098  
CC represent oligonucleotide sequences and primers which are used in the  
CC exemplification of the present invention. N.B. The present invention  
CC specifically claims all the polynucleotide sequences given in the  
CC sequence listing of the present specification, however the sequence  
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
CC for SEQ ID NO:4455 to 4464  
XX  
SQ Sequence 266 AA;  
XX  
QY Query Match 38.6%; Score 22; DB 4; Length 266;  
Best Local Similarity 36.4%; Pred. No. 6.2e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
19 FXXXXXGXXF 29  
147 FKTRSGVYTF 157  
Db  
XX  
RESULT 14  
ABO69527 ID ABO69527 standard; protein; 268 AA.  
XX  
AC ABO69527;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #1702.  
XX  
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
XX  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
DR WPI; 2003-615309/58.  
XX  
N-PSDB; ABD03098.  
XX  
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 18273; 455pp; English.  
XX  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using biochip technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at

CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 268 AA;  
XX  
QY Query Match 38.6%; Score 22; DB 7; Length 268;  
Best Local Similarity 36.4%; Pred. No. 6.3e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
19 FXXXXXGXXF 29  
41 FLAVATGTYAF 51  
Db  
XX  
RESULT 15  
ABP39297 ID ABP39297 standard; protein; 292 AA.  
XX  
AC ABP39297;  
XX  
DT 24-JUL-2002 (first entry)  
XX  
DE Staphylococcus epidermidis ORF amino acid sequence SRQ ID NO:4142.  
XX  
KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
XX antibacterial; gene therapy.  
XX  
OS Staphylococcus epidermidis.  
XX  
PN US6380370-B1.  
XX  
PD 30-APR-2002.  
XX  
PF 13-AUG-1998; 98US-00134001.  
XX  
PR 14-AUG-1997; 97US-0055779P.  
XX  
PR 08-NOV-1997; 97US-0064964P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Doucette-Stamm LA, Bush D;  
XX  
DR WPI; 2002-381255/41.  
XX  
N-PSDB; ABN91842.  
XX  
PT Novel isolated nucleic acid encoding a Staphylococcus epidermis  
XX polypeptide, useful for diagnosing and treating bacterial infections.  
XX  
PS Disclosure; SEQ ID NO 4142; 267pp; English.  
XX  
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences  
XX given in ABP35124 to ABP37960. The *S. epidermidis* sequences have  
XX antibacterial activity and can be used in gene therapy. The sequences can  
XX also be used in the diagnosis and treatment of bacterial infections,  
XX particularly *S. epidermidis* infections. The sequences can be used to  
XX screen for compounds able to interfere with the *S. epidermidis* life cycle  
XX or inhibit *S. epidermidis* infection. N.B. The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from the USPTO web site  
XX  
SQ Sequence 292 AA;  
XX  
QY Query Match 38.6%; Score 22; DB 5; Length 292;  
Best Local Similarity 36.4%; Pred. No. 6.3e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
19 FXXXXXGXXF 29  
173 FKTRSGVYTF 163  
Db  
XX  
Search completed: December 25, 2004, 08:30:35  
Job time : 49.6487 secs



: Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:08:11 ; Search time 7.9462 Seconds  
(without alignments) 375.364 Million cell updates/sec

Title: US-10-621-787-5  
Perfect score: 57  
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXGXXXXXX 31

Scoring table: BIOSUM62  
Gapop 10.0 , Gapect 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 79: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	38.6	250	2	E83651
2	22	38.6	302	2	D82112
3	22	38.6	421	2	A99309
4	22	38.6	423	2	H90490
5	22	38.6	441	2	T19306
6	22	38.6	868	2	T02635
7	22	38.6	3705	2	AD0123
8	22	38.6	4976	2	T14165
9	21	36.8	156	2	E72052
10	21	36.8	156	2	A86573
11	21	36.8	210	2	F64609
12	21	36.8	210	2	B71906
13	21	36.8	212	2	H81344
14	21	36.8	231	2	B83060
15	21	36.8	238	2	I64038
16	21	36.8	269	2	T49293
17	21	36.8	277	2	T06760
18	21	36.8	303	2	T35616
19	21	36.8	303	2	A12543
20	21	36.8	307	2	A71602
21	21	36.8	325	2	T31989
22	21	36.8	420	2	AG3057
23	21	36.8	420	2	G98238
24	21	36.8	450	2	B96561
25	21	36.8	508	2	T04605
26	21	36.8	516	2	C82946
27	21	36.8	555	2	T01391
28	21	36.8	653	2	T03319
29	21	36.8	669	2	T24571

30	21	36.8	704	2	T13665	NADH2 dehydrogenas
31	21	36.8	724	2	AB0531	lysine decarboxyla
32	21	36.8	724	2	A42953	nitrous oxide redu
33	21	36.8	724	2	D96949	secreted protein c
34	21	36.8	779	2	T44659	nitrous oxide redu
35	21	36.8	826	2	E81706	conserved hypochet
36	21	36.8	885	2	AC2444	hypochetrical prote
37	21	36.8	975	2	G82209	GDDR family prote
38	21	36.8	1138	2	G71554	probable transmemb
39	21	36.8	1316	2	B86240	protein F20824.12
40	21	36.8	1417	2	H90670	probable invasin l
41	21	36.8	1417	2	D85521	probable adhesin e
42	20	35.1	40	2	G81731	hypochetrical prote
43	20	35.1	119	2	A53257	H+-transporting AT
44	20	35.1	127	2	H64012	sufi protein homol
45	20	35.1	129	2	B90175	NADH dehydrogenase
46	20	35.1	171	2	C81280	probable periplasm
47	20	35.1	173	2	I39569	hypochetrical prote
48	20	35.1	181	2	G75071	phenylacrylic acid
49	20	35.1	181	2	A71094	probable 3-octapre
50	20	35.1	187	2	AF2009	hypochetrical prote
51	20	35.1	219	2	F97083	protein of phospho
52	20	35.1	240	2	H69535	acetylglutamate ki
53	20	35.1	242	2	I78931	arylamine N-acetyl
54	20	35.1	246	2	T13747	hypochetrical prote
55	20	35.1	260	2	D24584	spectinomycin resi
56	20	35.1	267	2	B69752	beta-lactamase hom
57	20	35.1	282	2	G72369	hemk protein - The
58	20	35.1	290	2	I78930	arylamine N-acetyl
59	20	35.1	290	2	B61267	arylamine N-acetyl
60	20	35.1	290	2	I58425	llypolysaccharide
61	20	35.1	297	2	F64206	oligopeptide trans
62	20	35.1	298	2	AC3617	hypochetrical prote
63	20	35.1	299	2	H64354	conserved hypochet
64	20	35.1	299	2	AH0367	glutamate decarbox
65	20	35.1	304	2	AE3623	conserved hypochet
66	20	35.1	313	2	A95859	protein C50H11.6 (
67	20	35.1	317	2	F88986	gluconolactonase (
68	20	35.1	320	2	S28218	hypochetrical prote
69	20	35.1	323	2	T25459	hypochetrical prote
70	20	35.1	334	2	T19860	protein C50H11.4 (
71	20	35.1	353	2	A88987	probable potassium
72	20	35.1	366	2	T43361	hypochetrical prote
73	20	35.1	385	2	G64251	hypochetrical prote
74	20	35.1	385	2	S73488	hypochetrical prote
75	20	35.1	417	2	D90304	amino acid transpo
76	20	35.1	418	2	B75087	multidrug resistanc
77	20	35.1	424	2	E81358	glutamate-1-semial
78	20	35.1	426	2	AG0411	glutamate-1-semial
79	20	35.1	437	2	T32763	hypochetrical prote
80	20	35.1	438	2	A57720	hypochetrical prote
81	20	35.1	462	2	G95233	hypochetrical prote
82	20	35.1	462	2	A86717	hypochetrical prote
83	20	35.1	464	2	E83029	hypochetrical prote
84	20	35.1	465	2	C81751	hypochetrical prote
85	20	35.1	465	2	F71489	hypochetrical prote
86	20	35.1	466	1	B43332	glutamate decarbox
87	20	35.1	466	1	S24234	glutamate decarbox
88	20	35.1	466	1	E91178	glutamate decarbox
89	20	35.1	466	2	F86024	glutamate decarbox
90	20	35.1	466	2	G85726	glutamate decarbox
91	20	35.1	466	2	B90891	glutamate decarbox
92	20	35.1	467	2	D86583	glutamate decarbox
93	20	35.1	467	2	E72040	glutamate decarbox
94	20	35.1	472	2	A99098	glutamate decarbox
95	20	35.1	498	1	A60028	glutamate decarbox
96	20	35.1	498	1	VH1V34	glutamate decarbox
97	20	35.1	498	1	VH1V34	glutamate decarbox
98	20	35.1	498	1	VH1V34	glutamate decarbox
99	20	35.1	498	1	VH1V34	glutamate decarbox
100	20	35.1	498	1	VH1V34	glutamate decarbox

## ALIGNMENTS

## RESULT 1

hypothetical protein BH0013 [imported] - Bacillus halodurans (strain C-125)  
C/Species: Bacillus halodurans  
C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C/Accession: E83651  
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A/Reference number: A83650; MUID:20512582; PMID:11058132  
A/Accession: E83651  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-250 <STO>  
A/Cross-references: UNIPROT:Q9KGP4; GB:AE001507; GB:BA000004; NID:g10172612; PIDN:BA037  
A/Experimental source: strain C-125  
C/Genetics:  
A/Gene: BH0013

Query Match 38.6%; Score 22; DB 2; Length 250;  
Best Local Similarity 36.4%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 132 FTFSGAGSYDF 142

## RESULT 2

tyra protein VC2145 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
D82112  
C/Species: Vibrio cholerae  
C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C/Accession: D82112  
R/Heidelberg, J.F.; Bisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baser, S.; Qin, H.; Dragol, I.; Sellers, F.  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A/Reference number: A82035; MUID:20406833; PMID:10952301  
A/Accession: D82112  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-302 <HEI>  
A/Cross-references: UNIPROT:Q9KQ59; GB:AE004287; GB:AE003852; NID:g9656689; PIDN:AAF9528  
A/Experimental source: serogroup O1; strain N16961; biotype El Tor  
C/Genetics:  
A/Gene: VC2145  
A/Map position: 1  
C/Superfamily: Escherichia coli hypothetical protein b2431

Query Match 38.6%; Score 22; DB 2; Length 302;  
Best Local Similarity 36.4%; Pred. No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 277 FTKAVTGAATF 287

## RESULT 3

membrane transporter SSO1505 [imported] - Sulfolobus solfataricus  
C/Species: Sulfolobus solfataricus  
C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C/Accession: A99309  
R/She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139  
A/Accession: A99309  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-421 <KUR>  
A/Cross-references: UNIPROT:Q97Y43; GB:AE006641; NID:g13814734; PIDN:AAK41728.1; GSPDB:G  
A/Genes: SSO1505

Query Match 38.6%; Score 22; DB 2; Length 421;  
Best Local Similarity 36.4%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 338 FAATAATGYSF 348

## RESULT 4

membrane transporter SSO3079 [imported] - Sulfolobus solfataricus  
H90490  
C/Species: Sulfolobus solfataricus  
C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C/Accession: H90490  
R/She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A/Description: Sulfolobus solfataricus complete genome.  
A/Reference number: A99139  
A/Accession: H90490  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-423 <KUR>  
A/Cross-references: UNIPROT:Q97UD8; GB:AE006641; NID:g13816494; PIDN:AAK43183.1; GSPDB:G  
A/Genes: SSO3079

Query Match 38.6%; Score 22; DB 2; Length 423;  
Best Local Similarity 36.4%; Pred. No. 3.7e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 340 FAATAATGYSF 350

## RESULT 5

hypothetical protein C15C8.1 - Caenorhabditis elegans  
T19306  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T19306  
R/Sims, M.  
submitted to the EMBL Data Library, June 1996  
A/Reference number: Z19106  
A/Accession: T19306  
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA  
A/Residues: 1-441 <ML>  
A/Cross-references: UNIPROT:Q18017; EMBL:Z75527; PIDN:CAA99774.1; GSPDB:GN00023; CESP:C1;  
A/Experimental source: clone C15C8  
C/Genetics:  
A/Gene: CESP:C15C8.1  
A/Map position: 5  
A/Intons: 156/2; 205/3; 258/1

Query Match 38.6%; Score 22; DB 2; Length 441;  
Best Local Similarity 36.4%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29

Db 165 PAALISGSYHF 175

## RESULT 6

D2 protein homolog - slime mold (Dictyostelium discoideum) plasmid Dp5

C:Species: Dictyostelium discoideum

C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004

C:Accession: 102635

R:Rieben J.F., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyosawa, H.; Hugh

Genetics 148, 1117-1125, 1998

A:Title: Dictyostelium discoideum nuclear plasmid Dp5 is a chimera related to the Dp1

A:Reference number: 214684; PMID:98198836; PMID:9539429

A:Accession: 102635

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-868 <RIB>

A:Cross-references: UNIPROT:O60984; EMBL:AF000580; NID:g3068582; PIDN:AA014391.1; PID:g3

A:Experimental source: strain MS2162

C:Genetics:

A:Gene: d2

A:Genome: plasmid

A:Mobile element: plasmid Dp5

C:Superfamily: slime mold (Dictyostelium discoideum) plasmid Dp5 D2 protein homolog

Query Match 38.6%; Score 22; DB 2; Length 868;

Best Local Similarity 36.4%; Pred. No. 7.6e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYXF 29

Db 345 FTTYSKGRYSF 355

## RESULT 7

probable autotransporter protein yaph [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004

C:Accession: AD0123

R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; PMID:21470413; PMID:11586360

A:Accession: AD0123

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-3705 <KUR>

A:Cross-references: UNIPROT:Q8ZHA1; GB:AL590842; PIDN:CAC9847.1; PID:g15979073; GSPDB:C

C:Genetics:

A:Gene: yaph

Query Match 38.6%; Score 22; DB 2; Length 3705;

Best Local Similarity 36.4%; Pred. No. 3.3e+03;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYXF 29

Db 3553 FTASVGGYAF 3563

## RESULT 8

peptide synthetase homolog - Mycobacterium smegmatis

C:Species: Mycobacterium smegmatis

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: 114165

R:Yu, S.; Flier, E.; Jacobs Jr., W.R.

J. Bacteriol. 180, 4676-4685, 1998

A:Title: Analysis of the exochelin locus in mycobacterium smegmatis: biosynthesis genes

A:Reference number: 217898; PMID:98389687; PMID:9721311

A:Accession: 114165

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4976 <YUS>

A:Cross-references: UNIPROT:O87314; EMBL:AF027770; NID:g3560502; PID:g3560507; PIDN:AA08

C:Genetics:

A:Gene: fxbC

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:58-566/Domain: acetate-CoA ligase homology <ACL>

F:582-649/Domain: acyl carrier protein homology <ACP1>

F:1620-2062/Domain: acetate-CoA ligase homology <ACL1>

F:2078-2144/Domain: acyl carrier protein homology <ACP2>

F:3139-3591/Domain: acetate-CoA ligase homology <ACL2>

F:3608-3679/Domain: acyl carrier protein homology <ACP3>

F:4198-4620/Domain: acetate-CoA ligase homology <ACL3>

F:4637-4705/Domain: acyl carrier protein homology <ACP4>

F:614,3643,4669/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 38.6%; Score 22; DB 2; Length 4976;

Best Local Similarity 36.4%; Pred. No. 4.5e+03;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYXF 29

Db 1738 FAGTSGLYDF 1748

## RESULT 9

rRNA methylase - Chlamydia pneumoniae (strain CWL029)

C:Species: Chlamydia pneumoniae

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: E72052

R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lamnel, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; PMID:99206606; PMID:10192388

A:Accession: E72052

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-156 <ARN>

A:Cross-references: UNIPROT:Q9Z7P4; GB:AE001648; GB:AE001363; NID:g4376946; PIDN:AA018791;

C:Genetics:

A:Experimental source: strain CWL029

C:Superfamily: Chlamydia pneumoniae rRNA methylase

Query Match 36.8%; Score 21; DB 2; Length 156;

Best Local Similarity 36.4%; Pred. No. 2.4e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYXF 29

Db 88 FSLPSSGTYVF 98

## RESULT 10

rRNA methylase [imported] - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: A86573

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.

A:Reference number: A86491; PMID:20330349; PMID:10871362

A:Accession: A86573

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-156 <STO>

A:Cross-references: UNIPROT:Q9Z7P4; GB:BA000008; NID:g8979032; PIDN:BA098867.1; GSPDB:GNI

A:Experimental source: strain J138

C/Genetics:  
A/Gene: spou 2  
C/Superfamily: Chlamydomphila pneumoniae rRNA methylase

Query Match 36.8%; Score 21; DB 2; Length 156;  
Best Local Similarity 36.4%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 88 FSLPSSGTYVF 98

# RESULT 11

conserved hypothetical integral membrane protein HPO718 - Helicobacter pylori (strain 26  
C/Species: Helicobacter pylori  
C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C/Accession: F64609

R/Tomb, J.F.; White, O.; Kervilavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.; Nature 388, 539-547, 1997  
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A/Reference number: A64520; MUID:97394467; PMID:9252185  
A/Accession: F64609

A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-210 <DOM>  
A/Cross-references: UNIPROT:O25420; GB:AE000585; GB:AE000511; NID:92313845; PIDD:AA00777

C/Superfamily: Escherichia coli hypothetical 23k protein (sbm-fba intergenic region)

Query Match 36.8%; Score 21; DB 2; Length 210;  
Best Local Similarity 36.4%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 75 FGALFTGFYAF 85

# RESULT 12

B71906  
C/Species: Helicobacter pylori (strain J99)  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C/Accession: B71906

R/Alt, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jlang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A/Reference number: A71800; MUID:99120557; PMID:9923682  
A/Accession: B71906

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-210 <ARN>

A/Cross-references: UNIPROT:Q9ZIC4; GB:AE001497; GB:AE001439; NID:94155199; PIDD:AA00622  
A/Experimental source: strain J99  
C/Genetics:

A/Gene: jhp0656  
C/Superfamily: Escherichia coli hypothetical 23k protein (sbm-fba intergenic region)

Query Match 36.8%; Score 21; DB 2; Length 210;  
Best Local Similarity 36.4%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 75 FGAVTGFYAF 85

# RESULT 13

H81344  
C/Species: Campylobacter jejuni  
C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C/Accession: H81344

R/Parhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell, Nature 403, 665-668, 2000  
A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A/Reference number: A81250; MUID:20150912; PMID:10688204  
A/Accession: H81344

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-212 <PAR>

A/Cross-references: UNIPROT:Q9PPH4; GB:AL139076; GB:AL111168; NID:96968128; PIDD:CA57300;  
A/Experimental source: serotype O2, strain NCTC 11168  
C/Genetics:

Query Match 36.8%; Score 21; DB 2; Length 212;  
Best Local Similarity 36.4%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 137 FSGKAVGYVSF 147

# RESULT 14

B83060  
C/Species: Pseudomonas aeruginosa  
C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C/Accession: B83060

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathoc  
A/Reference number: A82950; MUID:20437337; PMID:10984043  
A/Accession: B83060

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-231 <STO>

A/Cross-references: UNIPROT:Q9HVB0; GB:AE004882; GB:AE004091; NID:9950939; PIDD:AA00807;  
A/Experimental source: strain PA01  
C/Genetics:

A/Gene: PA4685

Query Match 36.8%; Score 21; DB 2; Length 231;  
Best Local Similarity 36.4%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29  
DB 146 FASEDSGVYRF 156

# RESULT 15

I64038  
C/Species: Haemophilus influenzae  
C/Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
C/Accession: I64038

R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kervilavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995  
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A/Reference number: A64000; MUID:95350630; PMID:7542800  
A/Accession: I64038

A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-238 <TTGR>  
 A:Cross-references: UNIPROT:P44278; GB:U32835; GB:I42023; NID:g1574459; PIDN:AAC23269.1

Query Match 36.8%; Score 21; DB 2; Length 238;  
 Best local similarity 36.4%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGYYX 29  
 | | | | |  
 Db 122 PALAGMGSYYP 132

Search completed: December 25, 2004, 08:39:40  
 Job time : 12.9462 secs

This Page Blank (upto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 25, 2004, 06:42:17 ; Search time 46.0095 Seconds

(without alignments)  
387.672 Million cell updates/sec

Title: US-10-621-787-5

Perfect score: 57  
Sequence: 1 PXXXXXXXXXXXXXXXXXXXXXGXFXFX 31Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	42.1	1020	2 Q89DX5	Q89dx5 bradyrhizob
2	23	40.4	352	2 Q7RAT5	Q7rat5 plasmidium
3	23	40.4	352	2 Q9DED3	Q9ded3 oncorhynch
4	23	40.4	483	1 PRTA_PHOAZ	P82115 photorhabd
5	23	40.4	483	2 Q84F72	Q84f72 photorhabd
6	23	40.4	483	2 AAS19409	Aas19409 photorhab
7	23	40.4	618	2 Q92LIN9	Q92lin9 rhizobium m
8	22	38.6	144	2 Q8BRU1	Q8bruj1 mus musculu
9	22	38.6	180	2 Q6N9E2	Q6n9e2 rhodospheudo
10	22	38.6	180	2 CAE27048	CaE27048 rhodospheu
11	22	38.6	188	2 Q8NBU8	Q8nbu8 homo sapien
12	22	38.6	215	1 YAO9_HUMAN	O9y3b3 homo sapien
13	22	38.6	217	2 Q6NWX3	Q6nwx3 brachydanio
14	22	38.6	217	2 AAH67383	AAh67383 brachydan
15	22	38.6	219	2 Q6P2V7	Q6p2v7 xenopus tro
16	22	38.6	219	2 AAH64279	AAh64279 xenopus t
17	22	38.6	224	2 Q96K51	Q96k51 homo sapien
18	22	38.6	224	2 Q848M1	Q848m1 uncultured
19	22	38.6	239	2 Q848M1	Q848m1 uncultured
20	22	38.6	250	2 Q9KGP4	Q9kgp4 bacillus ha
21	22	38.6	272	2 Q6PWJ1	Q6pwj1 ralistonia s
22	22	38.6	272	2 AA87586	AA87586 ralistonia s
23	22	38.6	290	2 Q8CSQ5	Q8csq5 staphylococ
24	22	38.6	296	2 Q9CNH6	Q9cnh6 pasteurella
25	22	38.6	298	2 Q6LTV2	Q6ltv2 photobacter
26	22	38.6	298	2 Q8G19273	Q8g19273 photobact
27	22	38.6	299	2 Q6D8U0	Q6d8u0 erwinia car
28	22	38.6	300	2 Q7MIM2	Q7mim2 vibrio vuln
29	22	38.6	300	2 Q8DMJ3	Q8dmj3 vibrio para
30	22	38.6	300	2 Q8DBA0	Q8dba0 vibrio vuln
31	22	38.6	302	2 Q9KQ59	Q9kq59 vibrio chol

32	22	38.6	323	2 Q8PXN5	Q8pxn5 methanosarc
33	22	38.6	325	2 Q8PSW8	Q8psw8 methanosarc
34	22	38.6	355	2 Q966G5	Q966g5 caenorhabdi
35	22	38.6	384	2 Q8EJZ4	Q8ejz4 shewanella
36	22	38.6	398	2 Q744P3	Q744p3 mycobacteri
37	22	38.6	398	2 AAS02705	Aas02705 mycobacte
38	22	38.6	404	2 Q8ULU7	Q8ulu7 pyrococcus
39	22	38.6	404	2 Q6UTJ2	Q6utj2 homo sapien
40	22	38.6	404	2 AAQ97430	AAq97430 homo sapi
41	22	38.6	404	2 AAQ97433	AAq97433 homo sapi
42	22	38.6	409	1 RUN3_MOUSE	Run3 mouse
43	22	38.6	409	2 Q91ZK1	Q91zk1 ratius norv
44	22	38.6	421	2 Q97Y43	Q97y43 sulfolobus
45	22	38.6	422	2 Q7ODI2	Q7odi2 anopheles g
46	22	38.6	423	2 Q97UD8	Q97ud8 sulfolobus
47	22	38.6	440	2 Q8XPA4	Q8xpa4 clostridium
48	22	38.6	441	2 Q18017	Q18017 caenorhabdi
49	22	38.6	480	1 PRTA_PHOLU	P82115 photorhabd
50	22	38.6	486	1 PRTA_PHOLU	Q84f70 photorhabd
51	22	38.6	519	2 Q7Q448	Q7q448 anopheles g
52	22	38.6	519	2 Q86GR1	Q86gr1 anopheles g
53	22	38.6	529	2 Q9F6F7	Q9f6f7 arabidopsis
54	22	38.6	608	2 Q7Q5X0	Q7q5x0 giardia lam
55	22	38.6	661	2 Q7QW29	Q7qw29 giardia lam
56	22	38.6	679	2 Q8XW21	Q8xw21 ralistonia s
57	22	38.6	743	2 Q89WU9	Q89wu9 bradyrhizob
58	22	38.6	763	2 Q6KGV2	Q6kgv2 bacterioph
59	22	38.6	787	2 Q9GSD4	Q9gsd4 plasmidium
60	22	38.6	859	2 Q7PDL4	Q7pdl4 plasmidium
61	22	38.6	868	2 Q6Q984	Q6q984 dictyosteli
62	22	38.6	973	2 Q7PJ63	Q7pj63 anopheles g
63	22	38.6	1098	2 Q8A049	Q8a049 bacteroides
64	22	38.6	1127	2 Q8L3A4	Q8l3a4 bacillus st
65	22	38.6	1457	2 Q9F289	Q9f289 yersinia pe
66	22	38.6	1677	2 Q9AS94	Q9as94 oryza sativ
67	22	38.6	1913	2 Q9GRJ6	Q9grj6 leucophaea
68	22	38.6	2215	2 Q7WBN0	Q7wbno bordetella
69	22	38.6	3346	2 Q7KNS4	Q7wns4 bordetella
70	22	38.6	3705	2 Q9F285	Q9f285 yersinia pe
71	22	38.6	3705	2 Q8ZHA1	Q8zha1 yersinia pe
72	22	38.6	3710	2 Q74QP7	Q74qp7 yersinia pe
73	22	38.6	3710	2 Q8CZU2	Q8czu2 yersinia pe
74	22	38.6	3710	2 AAS63576	AA63576 yersinia
75	22	38.6	4976	2 Q87314	Q87314 mycobacteri
76	22	38.6	5020	2 Q8B9W3	Q8b9w3 shewanella
77	22	38.6	6939	2 Q7R1S4	Q7rie4 giardia lam
78	22	36.8	45	2 Q8F630	Q8f630 leptospira
79	22	36.8	109	2 Q9NHK7	Q9nhk7 plasmidium
80	22	36.8	109	2 Q9NHL5	Q9nhl5 plasmidium
81	22	36.8	113	2 Q86TZ2	Q86tz2 dictyosteli
82	22	36.8	125	2 Q6MLU8	Q6mlu8 bdellovibri
83	22	36.8	125	2 CAE79758	CaE79758 bdellovib
84	22	36.8	135	2 Q8FNA7	Q8fna7 escherichia
85	22	36.8	136	2 Q6KGS1	Q6kgs1 macaca mula
86	22	36.8	136	2 AAQ02794	AAq02794 macaca mu
87	22	36.8	139	2 Q7Z239	Q7z239 plasmidium
88	22	36.8	154	2 Q9LSR0	Q9lsr0 salmonella
89	22	36.8	154	2 Q935K2	Q935k2 salmonella
90	22	36.8	156	2 Q9Z7P4	Q9z7p4 chlamydia p
91	22	36.8	159	2 Q8C1Z9	Q8c1z9 mus musculu
92	22	36.8	169	2 Q8E085	Q8e085 streptococc
93	22	36.8	169	2 Q8ESW2	Q8esw2 streptococc
94	22	36.8	175	1 Y995_THETN	Y995 thetinae
95	22	36.8	187	1 Y177_METWA	Y177 metanoeac
96	22	36.8	192	1 Q16605	Q16605 caenorhabdi
97	22	36.8	195	1 YG14_METAC	Yg14 methanosarc
98	22	36.8	196	1 P24_CRIGR	P24 cricetulus
99	22	36.8	196	1 YB05_THRAC	Yb05 thermoplasm
100	22	36.8	201	1 P24_HUMAN	P24 human

## ALIGNMENTS



```

RESULT 1
ID 089DX5 PRELIMINARY; PRT; 1020 AA.
AC 089DX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ACB/ACrd/ACrf family protein.
OS OrderedLocustNames=bl17312;
GN Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobiium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USD110;
RA MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamideawa K., Uchiyumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Irituchi M., Kawashima K.,
RA Kobara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USD110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005961; BAC52577.1; -.
DR HSSP; P31224; 11WG.
DR InterPro; IPR001036; Acrlflin_res.
DR Pfam; PF00873; ACR_Etan; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
KM Complete proteome.
SQ SEQUENCE 1020 AA; 109742 MW; F593D2PF5D6FDB9D CRC64;

Query Match 42.1%; Score 24; DB 2; Length 1020;
Best Local Similarity 36.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXXYP 29
Db 453 FAASSAGEYTP 463

RESULT 2
ID 07RAT5 PRELIMINARY; PRT; 214 AA.
AC 07RAT5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transmembrane protein tmp21 precursor.
GN Name=PY06414;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RA PubMed=12368865;
RA Carlton J.M., Angiolini S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kogack D.S., Shumway M.F., Biddlewell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M.,
RA Florens L., Yates P.R., Ritt Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B.,
RA van Lin L.H., Jense C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Cauceli D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

```

```

CC preliminary data.
DR EMBL; AAB01002169; EAA16629.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008320; F:protein carrier activity; IEA.
DR GO; GO:0006886; P:intracellular protein transport; IEA.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR InterPro; IPR009038; GOLD.
DR Pfam; PF01105; EMP24_GP25L; 1.
DR PROSITE; PSS0866; GOLD; 1.
SQ SEQUENCE 214 AA; 24900 MW; 21A20BFA65BCAF59 CRC64;

Query Match 40.4%; Score 23; DB 2; Length 214;
Best Local Similarity 36.4%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXXYP 29
Db 96 FTFYSGSYSP 106

RESULT 3
ID 09DED3 PRELIMINARY; PRT; 352 AA.
AC 09DED3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Inhibin.
GN Inhibin.
OS Oncomyrnchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Teleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Tada T., Endo M., Hirono I., Takashima F., Aoki T.;
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the TGF-beta family.
DR EMBL; AB044566; BAB19272.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008083; F:growth factor activity; IEA.
DR GO; GO:0008151; P:cell growth and/or maintenance; IEA.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFp.
DR Pfam; PF00019; TGF_beta; 1.
DR PRINTS; PR00669; INHIBIN.
DR PRODOM; PD000357; TGFp; 1.
DR SMART; SM00204; TGFp; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KM Growth factor.
SQ SEQUENCE 352 AA; 39711 MW; ED9CC8C6860F912ED CRC64;

Query Match 40.4%; Score 23; DB 2; Length 352;
Best Local Similarity 36.4%; Pred. No. 9.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXXYP 29
Db 326 FTTSDGYSYP 336

RESULT 4
ID PRTA_PHOAZ STANDARD; PRT; 483 AA.
AC P82115;
DT 05-JUN-2004 (Rel. 44, Created)
DT 05-JUN-2004 (Rel. 44, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Secreted alkaline metalloproteinase (EC 3.4.24.-).
GN Name=PRTA;
OS Photobacterium sp. (strain Az29).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

OC Enterobacteriaceae; Photorhabdus.  
 OX NCBI\_TaxID=229779;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mendes-Cabral C., Montiel R., Simoes N.;  
 RT "Cloning and sequencing of the gene for a metalloprotease with an ABC  
 RT transport system: a protease secreted by Photorhabdus sp. A229  
 RT involved in the inhibition of insect antibacterial peptides.";  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 155-162, 166-197 AND 264-277, AND SUBCELLULAR LOCATION.  
 RA Cherqui A., Mendes-Cabral C., Simoes N.;  
 RT "Two exoproteases from Photorhabdus luminescens.";  
 RL Submitted (SEP-1999) to Swiss-Prote.  
 CC -1- FUNCTION: Involved in the inhibition of insect antibacterial  
 CC peptides.  
 CC -1- COFACTOR: Binds 1 zinc ion and a number of calcium ions per  
 CC subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to peptidase family M10B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AY531111; AAS19409.1; -  
 DR InterPro: IPR001343; Hemolysin\_Ca\_bind.  
 DR InterPro: IPR006025; Pept\_M\_Zn\_BS.  
 DR PROSITE: PS00330; HEMOLYSIN\_CALCITUM, 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE, 1.  
 KM Calcium-binding; Direct protein sequencing; Hydrolyase; Metal-binding;  
 KM Metalloprotease; Repeat; Zinc.  
 FT REPEAT 343 351 GKGXD 1.  
 FT REPEAT 352 360 GKGXD 2.  
 FT REPEAT 361 369 GKGXD 3.  
 FT REPEAT 370 378 GKGXD 4.  
 FT REPEAT 379 387 GKGXD 5.  
 FT METAL 184 184 Zinc (catalytic) (By similarity).  
 FT ACT SITE 185 185 By similarity.  
 FT METAL 188 188 Zinc (catalytic) (By similarity).  
 FT METAL 194 194 Zinc (catalytic) (By similarity).  
 FT METAL 263 263 Calcium 1 (via carbonyl oxygen) (By  
 similarity).  
 FT METAL 266 266 Calcium 1 (By similarity).  
 FT METAL 295 295 Calcium 1 (By similarity).  
 FT METAL 297 297 Calcium 1 (via carbonyl oxygen) (By  
 similarity).  
 FT METAL 298 298 Calcium 2 (via carbonyl oxygen) (By  
 similarity).  
 FT METAL 300 300 Calcium 1 and 2 (By similarity).  
 FT METAL 337 337 Calcium 2 (via carbonyl oxygen) (By  
 similarity).  
 FT METAL 339 339 Calcium 2 (By similarity).  
 SQ SEQUENCE 483 AA; 53482 MW; F402483BD50293A1 CRC64;  
 Query Match 40.4%; Score 23; DB 1; Length 483;  
 Best Local Similarity 36.4%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 19 FXXXXXGXYYF 29  
 Db 135 FDASLTGSYAF 145  
 RESULT 5  
 ID 084F72 PRELIMINARY; PRT; 483 AA.  
 AC 084F72;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Secreted alkaline metalloprotease.  
 GN Name=PRTA;  
 OS Photorhabdus luminescens (Xenorhabdus luminescens).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Photorhabdus.  
 OX NCBI\_TaxID=29488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K122;  
 RA Waterfield N.;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY230749; AAO39136.1; -  
 DR HSSP: O69771; 1G9K.  
 DR MEROPS: M10.060; -  
 DR GO: GO:0005509; F:calcium ion binding; IEA.  
 DR GO: GO:0008237; F:metalloprotease activity; IEA.  
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro: IPR001343; Hemolysin\_Ca\_bind.  
 DR InterPro: IPR006026; Peptidase\_N.  
 DR InterPro: IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro: IPR011049; Seratalyn\_Like\_C.  
 DR Pfam: PF00353; HemolysinCabindf. 2.  
 DR PRINTS: PR00313; CABNDGRPT.  
 DR SMART: SMO0235; ZINC, 1.  
 DR PROSITE: PS00330; HEMOLYSIN\_CALCITUM, 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN.1.  
 SQ SEQUENCE 483 AA; 53354 MW; 9D95895B915DE09C CRC64;

Query Match 40.4%; Score 23; DB 2; Length 483;  
 Best Local Similarity 36.4%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 19 FXXXXXGXYYF 29  
 Db 135 FDASLTGSYAF 145

RESULT 6  
 ID AAS19409 PRELIMINARY; PRT; 483 AA.  
 AC AAS19409;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE Secreted alkaline metalloprotease.  
 GN PRTA.  
 OS Photorhabdus sp. A229.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Photorhabdus.  
 OX NCBI\_TaxID=229779;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A229;  
 RA Cabral C.M., Montiel R., Simoes N.;  
 RT "Cloning and sequencing of the gene for a metalloprotease with an ABC  
 RT transport system: a protease secreted by Photorhabdus sp. A229  
 RT involved in the inhibition of insect antibacterial peptides.";  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY531111; AAS19409.1; -  
 DR EMBL: AY531111; AAS19409.1; -  
 KM Metalloprotease; Protease.  
 SQ SEQUENCE 483 AA; 53482 MW; F402483BD50293A1 CRC64;

Query Match 40.4%; Score 23; DB 2; Length 483;  
 Best Local Similarity 36.4%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 19 FXXXXXGXYYF 29  
 Db 135 FDASLTGSYAF 145

```

RESULT 7
Q92LNG PRELIMINARY; PRT; 618 AA.
ID Q92LNG
AC Q92LNG
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein SMC03097.
GN OrderedLocustNames=R02996; ORFNames=SMC03097;
OS Rhizobium meliloti (Sinorhizobium meliloti);
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OC NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=2136623; PubMed=11474104;
RA Galibert F., Finan T.W., Long S.R., Puehler A., Abola P., Ampe F.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capella D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federapfel N.A., Fisher R.P.,
RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kaiman S., Keating D.H., Kise E., Komp C., Lelaure V.,
RA Masny D., Palm C., Beck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA Ramsperger U., Surzycki R., Thebaud P., Vandenbol M.,
RA Vorholter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batur J.;
RT "the composite genome of the legume symbiont Sinorhizobium meliloti.";
RL Science 293:668-672(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21366507; PubMed=11481430;
RA Capella D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batur J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kise E., Lelaure V., Masny D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
RN [3]
RX EMBL, AL591792; CAC47575.1;
DR InterPro; IPR00184; Bac surfAg D15.
DR InterPro; IPR010827; Surf Ag VNR.
DR InterPro; IPR001412; rRNA-synt_1.
DR Pfam; PF01103; Bac surface_Ag_1.
DR Pfam; PF07244; Surf_Ag_VNR_1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 618 AA; 65655 MW; E51276BB29B1FB78 CRC64;

Query Match 40.4%; Score 23; DB 2; Length 618;
Best Local Similarity 36.4%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Cy 19 FXXXXXGXYXF 29
Db 464 FEASASGYVAF 474

RESULT 8
Q8BPJ1 PRELIMINARY; PRT; 144 AA.
ID Q8BPJ1
AC Q8BPJ1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
DE library, clone:813032001 product:runt related transcription factor 3,
DE full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX MEDLINE=20493374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subcloning of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX MEDLINE=20330913; PubMed=11078661;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama Y., Nishi K., Kitsumai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaishiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Matshiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirokawa T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kaoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Miyata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohata N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi P., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK053910; BAC35587.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000040; AML1_Runt.
DR PRINTS; PR00967; ONCOGENEAML1.
SQ SEQUENCE 144 AA; 14731 MW; EC4789698ADB9D7 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 144;
Best Local Similarity 36.4%; Pred. No. 7.3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

Qy 19 FXXXXXGXYP 29  
Db 66 FYGASGSSTYGF 76

## RESULT 9

Q6N9E2 PRELIMINARY; PRT; 180 AA.  
ID Q6N9E2  
AC Q6N9E2;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DE Apr-1'-p processing enzyme family protein homolog.  
GN OrderedLocustNames=RP1607;  
OS Rhodopseudomonas palustris.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradythiobacteriaceae; Rhodopseudomonas.  
OX NCBI\_TaxID=1076;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CGA009 / ATCC BAA-98;  
RX PubMed=14704707; DOI=10.1038/nbt923;  
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,  
RA Land M.L., Pelletier D.A., Bobst C., Torres Y Torres J.L., Peres C.,  
RA Gibson J.L., Hanson T.E., Bobst C., Torres Y Torres J.L., Peres C.,  
RA Harrison F.H., Gibson J., Harwood C.S.;  
RT "Complete genome sequence of the metabolically versatile  
RT photosynthetic bacterium Rhodopseudomonas palustris.";  
RL Nat. Biotechnol. 22:55-61(2004).  
DR EMBL; BX572598; CAB27048.1; -  
DR InterPro; IPR002589; Alpp.  
DR Pfam; PF01661; Alpp; 1.  
DR SMART; SM00506; Alpp; 1.  
KM Complete proteome.  
SQ SEQUENCE 180 AA; 18435 MW; ABD79473EDDE394 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 180;  
Best Local Similarity 36.4%; Pred. No. 9e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYP 29  
Db 122 FSAISTGVYGF 132

## RESULT 10

CAB27048 PRELIMINARY; PRT; 180 AA.  
ID CAB27048  
AC CAB27048;  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DE Apr-1'-p processing enzyme family protein homolog.  
GN RP1607.  
OS Rhodopseudomonas palustris.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradythiobacteriaceae; Rhodopseudomonas.  
OX NCBI\_TaxID=1076;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CGA009 / ATCC BAA-98;  
RX PubMed=14704707;  
RA Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,  
RA Land M.L., Pelletier D.A., Bobst C., Torres Y Torres J.L., Peres C.,  
RA Gibson J.L., Hanson T.E., Bobst C., Torres Y Torres J.L., Peres C.,  
RA Harrison F.H., Gibson J., Harwood C.S.;  
RT "Complete genome sequence of the metabolically versatile  
RT photosynthetic bacterium Rhodopseudomonas palustris.";  
RL Nat. Biotechnol. 22:55-61(2004).  
DR EMBL; BX572598; CAB27048.1; -  
SQ SEQUENCE 180 AA; 18435 MW; ABD79473EDDE394 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 180;  
Best Local Similarity 36.4%; Pred. No. 9e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYP 29  
Db 122 FSAISTGVYGF 132

## RESULT 11

Q8NB08 PRELIMINARY; PRT; 188 AA.  
ID Q8NB08  
AC Q8NB08;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 25, Last sequence update)  
DE Hypothetical protein FLJ90737.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Isogai T., Oca T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,  
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,  
RA Masuho Y., Oho T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,  
RA Hattori A., Okumura K., Iwayanagi T., Minomiyu K.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK075218; BAC11479.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0008320; F:protein carrier activity; IEA.  
DR GO; GO:0006886; P:intracellular protein transport; IEA.  
DR InterPro; IPR000348; Emp24\_gp25L\_p24.  
DR InterPro; IPR009038; GOLD.  
DR Pfam; PF01105; EMP24\_gp25L; 1.  
DR PROSITE; PS50866; GOLD.  
SQ SEQUENCE 188 AA; 21233 MW; 15A10CC4BD553DB3 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 188;  
Best Local Similarity 36.4%; Pred. No. 9.4e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYP 29  
Db 98 FTASKNGTYGF 108

## RESULT 12

YA09 HUMAN STANDARD; PRT; 215 AA.  
ID YA09 HUMAN  
AC Q9Y3B3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein CGI-109 precursor.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20272150; PubMed=10810093;  
RX Lai C.-H., Chou C.-Y., Chang L.-Y., Liu C.-S., Lin W.-C.;  
RT "Identification of novel human genes evolutionarily conserved in  
RT Caenorhabditis elegans by comparative proteomics.";  
RL Genome Res. 10:703-713(2000).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic  
CC reticulum (by similarity).  
CC -1- SIMILARITY: Belongs to the EMP24/GP25L family.  
CC -1- SIMILARITY: Contains 1 GOLD domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.1db-sib.ch/announce/>  
CC or send an email to [license@1db-sib.ch](mailto:license@1db-sib.ch)).

DR EMBL/AF151667; AAD34104.1; -  
DR InterPro: IPR000348; Emp24\_gp25L\_p24.  
DR InterPro: IPR009038; GOLD.  
DR Pfam: PF01105; EMP24\_GP25L\_1.  
DR PROSITE; PS50866; GOLD; 1.  
KM Endoplasmic reticulum; Hypothetical protein; Signal; Transmembrane.  
FT SIGNAL 1 23 Potential.  
FT CHAIN 24 215 Hypothetical protein CG1-109.  
FT DOMAIN 24 176 Luminal (Potential).  
FT TRANSMEM 177 199 Potential.  
FT DOMAIN 200 215 Cytoplasmic (Potential).  
FT DOMAIN 38 120 GOLD.  
SQ SEQUENCE 215 AA; 24354 MW; 5C4D5BBE8054857B CRC64;

Query Match 38.6%; Score 22; DB 1; Length 215;  
Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

19 FXXXXXGXYYF 29  
90 FTSKNGTYKF 100

RESULT 13  
ID 06NMK3 PRELIMINARY; PRT; 217 AA.  
AC 06NMK3;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE Hypothetical protein zgc:76996.  
GN Name=zgc:76996;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_Taxid=7955;  
[1]  
SEQUENCE FROM N.A.

RP TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalka U., Smalins D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RX Strauberg R.;

RA TISSUE=Embryo;  
RA Strauberg R.;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL/BC067383; AAH67383.1; -  
DR InterPro: IPR000348; Emp24\_gp25L\_p24.  
DR InterPro: IPR009038; GOLD.  
DR Pfam: PF01105; EMP24\_GP25L\_1.  
DR PROSITE; PS50866; GOLD; 1.  
DR Hypothetical protein.  
SQ SEQUENCE 217 AA; 24555 MW; F885F671CDB86010 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 217;  
Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

19 FXXXXXGXYYF 29  
91 FSAARNGTYKF 101

RESULT 14  
ID AAH67383 PRELIMINARY; PRT; 217 AA.  
AC AAH67383;  
DT 24-MAY-2004 (TReMBLrel. 27, Created)  
DT 24-MAY-2004 (TReMBLrel. 27, Last sequence update)  
DT 24-MAY-2004 (TReMBLrel. 27, Last annotation update)  
DE Hypothetical protein zgc:76996.  
GN zgc:76996.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_Taxid=7955;  
[1]  
SEQUENCE FROM N.A.

RP TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalka U., Smalins D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RX Strauberg R.;

RA Strauberg R.;  
DR EMBL/BC067383; AAH67383.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 217 AA; 24555 MW; F885F671CDB86010 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 217;  
Best Local Similarity 36.4%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

19 FXXXXXGXYYF 29  
91 FSAARNGTYKF 101

## RESULT 15

06P2V7

PRELIMINARY: PRT: 219 AA.

AC 06P2V7

DT 05-JUL-2004 (TRENBLrel. 27, Created)

DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)

DE Hypothetical protein MGC76316.

GN

OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=8364;

RN

RP [1]

SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=22388257; PubMed=12477932;

RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stadleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Uedlin T.B., Toshlyuk S., Carninci P., Prange C.,

RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smalick D.E., Schnerch A., Schein J.E.,

RT "Generation and initial analysis of more than 15,000 full-length human

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RL

[2]

SEQUENCE FROM N.A.

RP TISSUE=Embryo;

RA Klein S., Strauberg R.,

RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC064279; AA064279.1; \_5P25L\_P24.

DR InterPro; IPR000348; EMP24\_GP25L\_P24.

DR Pfam; PF01105; EMP24\_GP25L; 1.

DR PROSITE; PS50866; GOLD; 1.

KM Hypothetical protein.

SQ SEQUENCE 219 AA; 25307 MW; 1D9BDF5626C4C14 CRC64;

Query Match 38.6%; Score 22; DB 2; Length 219;

Best Local Similarity 36.4%; Pred.No. 1.1e+03;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYR 29

Db 93 FTATRNGTYKF 103

Search completed: December 25, 2004, 08:38:30  
Job time : 60.0095 secs

1115 rye blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:12:38 ; Search time 13.4399 Seconds  
(without alignments)  
152.967 Million cell updates/sec

Title: US-10-621-787-5

Perfect score: 57  
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXXGXFXFX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	22	38.6	266 4 US-09-710-279-2066	Sequence 2066, Ap
2	22	38.6	268 4 US-09-252-991A-18273	Sequence 18273, A
3	22	38.6	292 3 US-09-134-001C-4142	Sequence 4142, Ap
4	22	38.6	305 4 US-09-328-352-6069	Sequence 6069, Ap
5	22	38.6	404 4 US-09-328-352-8182	Sequence 8182, Ap
6	21	36.8	156 4 US-09-198-452A-698	Sequence 698, App
7	21	36.8	201 2 US-08-801-740-6	Sequence 6, Appli
8	21	36.8	201 2 US-08-801-740-7	Sequence 7, Appli
9	21	36.8	201 3 US-08-801-740-6	Sequence 6, Appli
10	21	36.8	201 3 US-08-801-740-7	Sequence 7, Appli
11	21	36.8	252 4 US-09-252-991A-30813	Sequence 30813, A
12	21	36.8	450 4 US-09-489-039A-10052	Sequence 10052, A
13	21	36.8	1233 4 US-09-134-000C-4971	Sequence 4971, Ap
14	21	36.8	1416 4 US-09-071-035-404	Sequence 404, App
15	21	36.8	1448 4 US-09-071-035-402	Sequence 402, App
16	20	35.1	63 2 US-08-637-759B-501	Sequence 501, App
17	20	35.1	63 3 US-08-871-355A-501	Sequence 501, App
18	20	35.1	63 3 US-09-201-945-501	Sequence 501, App
19	20	35.1	86 4 US-09-178-093B-34	Sequence 34, Appl
20	20	35.1	110 4 US-09-270-767-44692	Sequence 44692, A
21	20	35.1	142 2 US-08-997-080-47	Sequence 47, Appl
22	20	35.1	142 2 US-08-997-362-47	Sequence 47, Appl
23	20	35.1	142 3 US-08-873-970-47	Sequence 47, Appl
24	20	35.1	142 3 US-09-095-855-47	Sequence 47, Appl
25	20	35.1	142 3 US-08-705-347A-47	Sequence 47, Appl
26	20	35.1	142 3 US-09-324-542-47	Sequence 47, Appl
27	20	35.1	142 4 US-09-205-426-47	Sequence 47, Appl

28	20	35.1	142 4 US-09-200-643-47	Sequence 47, Appl
29	20	35.1	143 3 US-08-685-808-2	Sequence 2, Appli
30	20	35.1	143 3 US-08-505-860C-2	Sequence 2, Appli
31	20	35.1	162 2 US-08-606-143-44	Sequence 44, Appl
32	20	35.1	226 4 US-09-270-767-44619	Sequence 44619, A
33	20	35.1	229 4 US-09-270-767-56999	Sequence 56999, A
34	20	35.1	312 4 US-09-543-681A-5883	Sequence 5883, Ap
35	20	35.1	329 4 US-09-248-796A-16263	Sequence 16263, A
36	20	35.1	336 3 US-08-987-691A-5	Sequence 5, Appli
37	20	35.1	408 4 US-09-252-991A-33131	Sequence 33131, A
38	20	35.1	429 4 US-09-922-501-10	Sequence 10, Appl
39	20	35.1	448 4 US-09-361-443-4	Sequence 4, Appli
40	20	35.1	462 4 US-09-583-110-3587	Sequence 3587, Ap
41	20	35.1	467 4 US-09-361-443-2	Sequence 2, Appli
42	20	35.1	469 4 US-09-198-452A-793	Sequence 793, App
43	20	35.1	492 4 US-09-252-991A-27853	Sequence 27853, A
44	20	35.1	498 3 US-09-232-468A-18	Sequence 18, Appl
45	20	35.1	498 3 US-09-232-468A-24	Sequence 24, Appl
46	20	35.1	498 3 US-08-686-968C-231	Sequence 231, App
47	20	35.1	498 4 US-09-784-984B-52	Sequence 52, Appl
48	20	35.1	498 4 US-09-784-984B-54	Sequence 54, Appl
49	20	35.1	527 4 US-09-103-331-4	Sequence 4, Appli
50	20	35.1	527 4 US-09-631-594-55	Sequence 55, Appl
51	20	35.1	546 2 US-09-067-351-1	Sequence 1, Appli
52	20	35.1	546 3 US-09-360-490-1	Sequence 10, Appl
53	20	35.1	646 3 US-09-336-643A-10	Sequence 10, Appl
54	20	35.1	681 4 US-09-248-796A-16367	Sequence 16367, A
55	20	35.1	755 4 US-09-270-767-45272	Sequence 45272, A
56	20	35.1	922 4 US-09-198-452A-15	Sequence 15, Appl
57	20	35.1	1312 4 US-09-489-039A-11544	Sequence 11544, A
58	20	35.1	1430 3 US-09-008-172-2	Sequence 2, Appli
59	20	35.1	1430 3 US-09-210-361-6	Sequence 6, Appli
60	20	35.1	1430 3 US-09-740-274-6	Sequence 6, Appli
61	20	35.1	3177 2 US-08-477-451-4	Sequence 4, Appli
62	19	33.3	32 3 US-08-525-539A-35	Sequence 35, Appl
63	19	33.3	59 4 US-09-621-976-7401	Sequence 7401, Ap
64	19	33.3	62 4 US-09-621-976-5072	Sequence 5072, Ap
65	19	33.3	62 4 US-09-621-976-5079	Sequence 5079, Ap
66	19	33.3	62 4 US-09-621-976-5080	Sequence 5080, Ap
67	19	33.3	62 4 US-09-621-976-5082	Sequence 5082, Ap
68	19	33.3	63 4 US-09-621-976-5078	Sequence 5078, Ap
69	19	33.3	74 4 US-09-621-976-7511	Sequence 7511, Ap
70	19	33.3	79 4 US-09-621-976-7507	Sequence 7507, Ap
71	19	33.3	91 4 US-09-248-796A-21815	Sequence 21815, A
72	19	33.3	99 4 US-09-621-976-7518	Sequence 7518, Ap
73	19	33.3	101 4 US-09-270-767-33955	Sequence 33955, A
74	19	33.3	101 4 US-09-270-767-54812	Sequence 54812, A
75	19	33.3	104 4 US-09-270-767-58325	Sequence 58325, A
76	19	33.3	107 4 US-09-344-624-7	Sequence 7, Appli
77	19	33.3	108 3 US-08-591-632-75	Sequence 75, Appl
78	19	33.3	108 3 US-08-591-632-85	Sequence 85, Appl
79	19	33.3	108 3 US-09-611-451-75	Sequence 75, Appl
80	19	33.3	108 3 US-09-611-451-85	Sequence 85, Appl
81	19	33.3	110 4 US-09-248-796A-21943	Sequence 21943, A
82	19	33.3	111 4 US-10-162-012-11	Sequence 11, Appl
83	19	33.3	115 4 US-08-956-171E-5217	Sequence 5217, Ap
84	19	33.3	115 4 US-08-781-986A-5217	Sequence 5217, Ap
85	19	33.3	122 1 US-08-249-013-4	Sequence 4, Appli
86	19	33.3	122 2 US-08-886-863-4	Sequence 4, Appli
87	19	33.3	122 3 US-09-175-229-4	Sequence 4, Appli
88	19	33.3	122 5 PCT-US95-06764-4	Sequence 4, Appli
89	19	33.3	123 4 US-09-621-976-6980	Sequence 680, Ap
90	19	33.3	126 4 US-09-710-279-2766	Sequence 2766, Ap
91	19	33.3	129 4 US-09-621-976-4973	Sequence 4763, Ap
92	19	33.3	130 3 US-09-485-316A-10	Sequence 10, Appl
93	19	33.3	131 4 US-09-107-532A-5741	Sequence 5741, Appl
94	19	33.3	133 4 US-09-583-110-5269	Sequence 5269, Ap
95	19	33.3	145 3 US-09-286-529-21	Sequence 21, Appl
96	19	33.3	153 4 US-09-270-767-4604	Sequence 4604, A
97	19	33.3	154 1 US-07-752-101A-69	Sequence 69, Appl
98	19	33.3	155 3 US-09-134-001C-5566	Sequence 5566, Ap
99	19	33.3	155 4 US-09-270-767-32485	Sequence 32485, A
100	19	33.3	155 4 US-09-270-767-47702	Sequence 47702, A

## ALIGNMENTS

## RESULT 1

US-09-710-279-2066  
; Sequence 2066, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P33480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2066  
; LENGTH: 266  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-2066

## Query Match

Best Local Similarity 38.6%; Score 22; DB 4; Length 266;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 147 FKTTRSGVYTF 157

## RESULT 2

US-09-252-991A-18273  
; Sequence 18273, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18273  
; LENGTH: 268  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18273

## Query Match

Best Local Similarity 38.6%; Score 22; DB 4; Length 268;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 41 FLAVATGTYAF 51

## RESULT 3

US-09-134-001C-4142  
; Sequence 4142, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4142  
; LENGTH: 292  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4142

## Query Match

Best Local Similarity 38.6%; Score 22; DB 3; Length 292;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 173 FKTTRSGVYTF 183

## RESULT 4

US-09-328-352-6069  
; Sequence 6069, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6069  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6069

## Query Match

Best Local Similarity 38.6%; Score 22; DB 4; Length 305;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYYF 29

DB 52 FTGTAQSTYTF 62

## RESULT 5

US-09-328-352-8182  
; Sequence 8182, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 8182  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-8182

## Query Match

Best Local Similarity 38.6%; Score 22; DB 4; Length 404;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGYXF 29

Db 242 FLKTTAGHYAF 252

## RESULT 6

US-09-198-452A-698

Sequence 698, Application US/09198452A

Patent No. 6559294

GENERAL INFORMATION:

APPLICANT: Grifflais, R.

TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

TITLE OF INVENTION: and treatment of infection

FILE REFERENCE: 9710-003-999

CURRENT APPLICATION NUMBER: US/09/198,452A

CURRENT FILING DATE: 1998-11-24

NUMBER OF SEQ ID NOS: 6849

SEQ ID NO 698

LENGTH: 156

TYPE: PRT

ORGANISM: Chlamydia pneumoniae

US-09-198-452A-698

Query Match 36.8%; Score 21; DB 4; Length 156;

Best Local Similarity 36.4%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGYXF 29

Db 88 FSLPSSGTYVF 98

## RESULT 7

US-08-801-740-6

Sequence 6, Application US/08801740

Patent No. 5869639

GENERAL INFORMATION:

APPLICANT: Goll, Surya K.

APPLICANT: Bandman, Olga

TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,740

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0189 US

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 6;

SEQUENCE CHARACTERISTICS:

LENGTH: 201 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Genbank

CLONE: 1212965

US-08-801-740-6

Query Match 36.8%; Score 21; DB 2; Length 201;

Best Local Similarity 36.4%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGYXF 29

Db 82 FAAMDGYTGF 92

## RESULT 8

US-08-801-740-7

Sequence 7, Application US/08801740

Patent No. 5869639

GENERAL INFORMATION:

APPLICANT: Goll, Surya K.

APPLICANT: Bandman, Olga

TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,740

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0189 US

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 7;

SEQUENCE CHARACTERISTICS:

LENGTH: 201 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Genbank

CLONE: 1213221

US-08-801-740-7

Query Match 36.8%; Score 21; DB 2; Length 201;

Best Local Similarity 36.4%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGYXF 29

Db 82 FAAMDGYTGF 92

```
RESULT 9
US-08-801-740-6
; Sequence 6, Application US/08801740
; Patent No. 6130325
; GENERAL INFORMATION:
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,740
; FILING DATE: 14-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0189 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1212965
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-801-740-6
Query Match 36.8%; Score 21; DB 3; Length 201;
Best Local Similarity 36.4%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGYXF 29
DB 82 FAAMDGTYKF 92
RESULT 10
US-08-801-740-7
; Sequence 7, Application US/08801740
; Patent No. 6130325
; GENERAL INFORMATION:
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```

```
ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,740
; FILING DATE: 14-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0189 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1213221
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-801-740-7
Query Match 36.8%; Score 21; DB 3; Length 201;
Best Local Similarity 36.4%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGYXF 29
DB 82 FAAMDGTYKF 92
RESULT 11
US-09-252-991A-30813
; Sequence 30813, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30813
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30813
Query Match 36.8%; Score 21; DB 4; Length 252;
Best Local Similarity 36.4%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGYXF 29
DB 167 FASEDSGVYRF 177
```

RESULT 12  
US-09-489-039A-10052  
Sequence 10052, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 10052  
LENGTH: 450  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10052

Query Match 36.8%; Score 21; DB 4; Length 450;  
Best Local Similarity 36.4%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGYXF 29  
DB 174 PARPDAGVAF 184

RESULT 13  
US-09-134-000C-4971  
Sequence 4971, Application US/09134000C  
Patent No. 6617136  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 4971  
LENGTH: 1233  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-4971

Query Match 36.8%; Score 21; DB 4; Length 1233;  
Best Local Similarity 36.4%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGYXF 29  
DB 990 FTTDEKQYGF 1000

RESULT 14  
US-09-071-035-404  
Sequence 404, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA

ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 404:  
SEQUENCE CHARACTERISTICS:  
INFORMATION FOR SEQ ID NO: 404:  
LENGTH: 1416 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-071-035-404

Query Match 36.8%; Score 21; DB 4; Length 1416;  
Best Local Similarity 36.4%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGYXF 29  
DB 959 FTTDEKQYGF 969

RESULT 15  
US-09-071-035-402  
Sequence 402, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 402:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1448 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-071-035-402

Query Match 36.8%; Score 21; DB 4; Length 1448;  
Best Local Similarity 36.4%; Pred. No. 1.6e+03;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXGXYP 29  
Db 986 FTTDEKGYSP 996

Search completed: December 25, 2004, 08:42:03  
Job time : 16.4399 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2004, 08:38:38 / Search time 54.5443 Seconds  
(without alignments)  
204.102 Million cell updates/sec

Title: US-10-621-787-5  
Perfect score: 1 FXXXXXXXXXXXXXXXXXXXXGXXFXFX 31  
Sequence: 57

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1595201 seqs, 359116952 residues

Total number of hits satisfying chosen parameters: 1595201

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubppaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	38.6	119	US-10-434-156-17	Sequence 17, Appl
2	38.6	136	US-10-434-156-22	Sequence 22, Appl
3	38.6	163	US-09-030-847-2	Sequence 2, Appl1
4	38.6	163	US-10-215-088-2	Sequence 2, Appl1
5	38.6	163	US-10-775-165-2	Sequence 2, Appl1
6	38.6	215	US-09-030-847-4	Sequence 4, Appl1
7	38.6	215	US-10-215-088-4	Sequence 4, Appl1
8	38.6	215	US-10-170-385-289	Sequence 289, Appl
9	38.6	215	US-10-775-165-4	Sequence 4, Appl1
10	38.6	219	US-10-276-774-2150	Sequence 67003, Ap
11	38.6	296	US-10-282-122A-67003	Sequence 67003, A
12	38.6	302	US-10-282-122A-77397	Sequence 77397, A
13	38.6	374	US-10-437-963-159816	Sequence 159816,

14	38.6	636	US-10-425-114-68352	Sequence 68352, A
15	38.6	636	US-10-425-115-274242	Sequence 274242, A
16	38.6	1421	US-10-424-599-275109	Sequence 275109, A
17	38.6	1678	US-10-437-963-126095	Sequence 126095, A
18	38.6	1730	US-10-424-599-275110	Sequence 275110, A
19	38.6	3705	US-10-282-122A-77944	Sequence 77944, A
20	36.8	53	US-10-029-386-29069	Sequence 29069, A
21	36.8	71	US-10-425-115-272794	Sequence 272794, A
22	36.8	71	US-10-425-115-272828	Sequence 272828, A
23	36.8	79	US-10-424-599-240256	Sequence 240256, A
24	36.8	83	US-10-424-599-168732	Sequence 168732, A
25	36.8	100	US-10-425-115-364312	Sequence 364312, A
26	36.8	128	US-10-466-242-51	Sequence 51, Appl1
27	36.8	131	US-10-767-701-42874	Sequence 42874, A
28	36.8	150	US-10-437-963-154089	Sequence 154089, A
29	36.8	156	US-10-289-762-698	Sequence 698, App
30	36.8	156	US-10-282-122A-54708	Sequence 54708, A
31	36.8	156	US-10-425-115-333955	Sequence 333955, A
32	36.8	158	US-10-425-115-333946	Sequence 333946, A
33	36.8	165	US-10-425-114-44317	Sequence 44317, A
34	36.8	201	US-09-833-245-2073	Sequence 2073, Ap
35	36.8	201	US-09-833-245-2075	Sequence 2075, Ap
36	36.8	201	US-09-833-245-2076	Sequence 2076, Ap
37	36.8	201	US-10-039-865-6	Sequence 6, Appl1
38	36.8	201	US-10-039-865-7	Sequence 7, Appl1
39	36.8	201	US-10-060-255-59	Sequence 59, Appl1
40	36.8	201	US-10-060-255-68	Sequence 68, Appl1
41	36.8	201	US-10-060-255-69	Sequence 69, Appl1
42	36.8	201	US-10-408-765A-717	Sequence 717, App
43	36.8	201	US-10-881-088-59	Sequence 59, Appl1
44	36.8	201	US-10-881-088-68	Sequence 68, Appl1
45	36.8	201	US-10-881-088-69	Sequence 69, Appl1
46	36.8	208	US-10-425-115-321090	Sequence 321090, A
47	36.8	210	US-10-335-977-5883	Sequence 5883, Ap
48	36.8	215	US-10-335-977-5884	Sequence 5884, Ap
49	36.8	215	US-10-335-977-5885	Sequence 5885, Ap
50	36.8	233	US-10-424-599-165427	Sequence 165427, A
51	36.8	238	US-10-282-122A-58562	Sequence 58562, A
52	36.8	277	US-10-425-114-43315	Sequence 43315, A
53	36.8	277	US-10-425-114-43316	Sequence 43316, A
54	36.8	300	US-09-769-736-86	Sequence 736, App
55	36.8	312	US-10-425-115-293160	Sequence 293160, A
56	36.8	336	US-10-363-616-321	Sequence 321, App
57	36.8	356	US-10-425-115-190446	Sequence 190446, A
58	36.8	356	US-10-425-115-293155	Sequence 293155, A
59	36.8	356	US-10-739-930-8219	Sequence 8219, Ap
60	36.8	376	US-10-425-114-55341	Sequence 55341, A
61	36.8	397	US-10-437-963-132936	Sequence 132936, A
62	36.8	408	US-10-781-014-462	Sequence 462, App
63	36.8	412	US-10-425-114-43315	Sequence 43315, A
64	36.8	420	US-10-369-493-11376	Sequence 11376, A
65	36.8	420	US-10-369-493-11778	Sequence 11778, A
66	36.8	420	US-10-369-493-14778	Sequence 14778, A
67	36.8	420	US-10-369-493-14932	Sequence 14932, A
68	36.8	421	US-10-369-493-15269	Sequence 15269, A
69	36.8	429	US-10-369-493-12300	Sequence 12300, A
70	36.8	429	US-10-369-493-15506	Sequence 15506, A
71	36.8	429	US-10-369-493-15878	Sequence 15878, A
72	36.8	429	US-10-369-493-16255	Sequence 16255, A
73	36.8	441	US-09-738-626-5934	Sequence 5934, Ap
74	36.8	466	US-10-424-599-260503	Sequence 260503, A
75	36.8	501	US-10-425-1107-69	Sequence 69, Appl1
76	36.8	508	US-10-342-224-42	Sequence 42, Appl1
77	36.8	525	US-10-424-599-192279	Sequence 192279, A
78	36.8	600	US-10-425-115-354835	Sequence 354835, A
79	36.8	661	US-10-369-493-10446	Sequence 10446, A
80	36.8	700	US-10-437-963-162308	Sequence 162308, A
81	36.8	803	US-10-437-963-154762	Sequence 154762, A
82	36.8	811	US-10-437-963-123314	Sequence 123314, A
83	36.8	845	US-10-437-963-163017	Sequence 163017, A
84	36.8	923	US-10-437-963-163015	Sequence 163015, A
85	36.8	1072	US-10-437-963-180020	Sequence 180020, A
86	36.8	1119	US-10-437-963-114130	Sequence 114130, A



87 21 36.8 1393 16 US-10-437-963-114133 Sequence 114133,  
88 21 36.8 1416 9 US-09-071-035-404 Sequence 404, App  
89 21 36.8 1416 14 US-10-206-576-404 Sequence 404, App  
90 21 36.8 1448 9 US-09-071-035-402 Sequence 402, App  
91 21 36.8 1448 14 US-10-206-576-402 Sequence 402, App  
92 21 36.8 1732 16 US-10-437-963-135872 Sequence 135872,  
93 20 35.1 20 14 US-10-123-101-84 Sequence 84, App1  
94 20 35.1 20 14 US-10-310-734-114 Sequence 114, App1  
95 20 35.1 55 14 US-10-029-386-28388 Sequence 28388, A  
96 20 35.1 57 15 US-10-424-599-205858 Sequence 205858,  
97 20 35.1 59 17 US-10-425-115-309502 Sequence 309502,  
98 20 35.1 63 15 US-10-424-599-220167 Sequence 220167,  
99 20 35.1 66 15 US-10-424-599-217628 Sequence 217628,  
100 20 35.1 71 16 US-10-437-963-155241 Sequence 155241,

## ALIGNMENTS

## RESULT 1

US-10-434-156-17  
; Sequence 17, Application US/10434156  
; Publication No. US20040019006A1  
; GENERAL INFORMATION:  
; APPLICANT: TORAY Industries Inc.  
; APPLICANT: Yoshihide HAYASHIZAKI  
; APPLICANT: Mamoru KAMITA  
; APPLICANT: Tomohiko SUZUKI  
; APPLICANT: Hiroe HIROKAWA  
; APPLICANT: Kaoru NAKAO  
; APPLICANT: Toshiaki TANAKA  
; TITLE OF INVENTION: No. US20040019006A1 genes relating to pain and use of the genes  
; FILE REFERENCE: 1254-0222P  
; CURRENT APPLICATION NUMBER: US/10/434,156  
; CURRENT FILING DATE: 2003-05-09  
; PRIOR FILING DATE: 2002-05-10  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 2.1  
; SEQ ID NO 17  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-434-156-17

Query Match 38.6%; Score 22; DB 15; Length 119;  
Best Local Similarity 36.4%; Pred. No. 5.1e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29  
DB 10 FTASRNGTYKF 20

## RESULT 2

US-10-434-156-22  
; Sequence 22, Application US/10434156  
; Publication No. US20040019006A1  
; GENERAL INFORMATION:  
; APPLICANT: TORAY Industries Inc.  
; APPLICANT: Yoshihide HAYASHIZAKI  
; APPLICANT: Mamoru KAMITA  
; APPLICANT: Tomohiko SUZUKI  
; APPLICANT: Hiroe HIROKAWA  
; APPLICANT: Kaoru NAKAO  
; APPLICANT: Toshiaki TANAKA  
; TITLE OF INVENTION: No. US20040019006A1 genes relating to pain and use of the genes  
; FILE REFERENCE: 1254-0222P  
; CURRENT APPLICATION NUMBER: US/10/434,156  
; CURRENT FILING DATE: 2003-05-09  
; PRIOR FILING DATE: 2002-05-10  
; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn version 2.1  
; SEQ ID NO 22  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-434-156-22

Query Match 38.6%; Score 22; DB 15; Length 136;  
Best Local Similarity 36.4%; Pred. No. 5.8e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29  
DB 10 FTASRNGTYKF 20

## RESULT 3

US-09-030-847-2  
; Sequence 2, Application US/09030847A  
; Patent No. US20020111472A1  
; GENERAL INFORMATION:  
; APPLICANT: NI, Jian  
; TITLE OF INVENTION: T1-R Ligand III  
; FILE REFERENCE: PF357  
; CURRENT APPLICATION NUMBER: US/09/030,847A  
; CURRENT FILING DATE: 1998-02-26  
; EARLIER APPLICATION NUMBER: 60/039,483  
; EARLIER FILING DATE: 1997-02-28  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-030-847-2

Query Match 38.6%; Score 22; DB 9; Length 163;  
Best Local Similarity 36.4%; Pred. No. 6.8e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29  
DB 89 FTASRNGTYKF 99

## RESULT 4

US-10-215-088-2  
; Sequence 2, Application US/10215088  
; Publication No. US20030069409A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: T1-R Ligand III  
; FILE REFERENCE: PF357C1  
; CURRENT APPLICATION NUMBER: US/10/215,088  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 09/030,847  
; PRIOR FILING DATE: 1998-02-26  
; PRIOR APPLICATION NUMBER: 60/039,483  
; PRIOR FILING DATE: 1997-02-28  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-215-088-2

Query Match 38.6%; Score 22; DB 14; Length 163;  
Best Local Similarity 36.4%; Pred. No. 6.8e+02;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29

Db 89 FTASKNGTYKF 99

## RESULT 5

US-10-775-165-2  
; Sequence 2, Application US/10775165  
; Publication No. US20040138443A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: TI-R Ligand III  
; FILE REFERENCE: PF357C1  
; CURRENT APPLICATION NUMBER: US/10/775,165  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: US/10/215,088  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 09/030,847  
; PRIOR FILING DATE: 1998-02-26  
; PRIOR APPLICATION NUMBER: 60/039,483  
; PRIOR FILING DATE: 1997-02-28  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Version 3.1  
; SEQ ID NO 2  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-775-165-2

## Query Match

Best Local Similarity 38.6%; Score 22; DB 16; Length 163;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYKF 29

Db 89 FTASKNGTYKF 99

## RESULT 6

US-09-030-847-4  
; Sequence 4, Application US/09030847A  
; Patent No. US2002011472A1  
; GENERAL INFORMATION:  
; APPLICANT: NI, Jian  
; TITLE OF INVENTION: TI-R Ligand III  
; FILE REFERENCE: PF357  
; CURRENT APPLICATION NUMBER: US/09/030,847A  
; CURRENT FILING DATE: 1998-02-26  
; EARLIER APPLICATION NUMBER: 60/039,483  
; EARLIER FILING DATE: 1997-02-28  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-030-847-4

## Query Match

Best Local Similarity 38.6%; Score 22; DB 9; Length 215;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYKF 29

Db 89 FTASKNGTYKF 99

## RESULT 7

US-10-215-088-4  
; Sequence 4, Application US/10215088  
; Publication No. US20030069409A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: TI-R Ligand III  
; FILE REFERENCE: PF357C1

; CURRENT APPLICATION NUMBER: US/10/215,088

; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 09/030,847  
; PRIOR FILING DATE: 1998-02-26  
; PRIOR APPLICATION NUMBER: 60/039,483  
; PRIOR FILING DATE: 1997-02-28  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Version 3.1  
; SEQ ID NO 4  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-215-088-4

## Query Match

Best Local Similarity 38.6%; Score 22; DB 14; Length 215;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYKF 29

Db 89 FTASKNGTYKF 99

## RESULT 8

US-10-170-385-289  
; Sequence 289, Application US/10170385  
; Publication No. US20030203372A1  
; GENERAL INFORMATION:  
; APPLICANT: Ward, Neil Raymond  
; APPLICANT: Mundy, Christopher Robert  
; APPLICANT: Kan, On  
; APPLICANT: Harris, Robert Alan  
; APPLICANT: White, Jonathan  
; APPLICANT: Binley, Katie Mary  
; APPLICANT: Rayner, William Nigel  
; APPLICANT: Naylor, Stuart  
; APPLICANT: Kingsman, Susan Mary  
; APPLICANT: Krige, David  
; TITLE OF INVENTION: ANALYSIS METHOD  
; FILE REFERENCE: S3268200100  
; CURRENT APPLICATION NUMBER: US/10/170,385  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: PCT/GB02/01662  
; PRIOR FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: PCT/GB01/05458  
; PRIOR FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 549  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 289  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-170-385-289

## Query Match

Best Local Similarity 38.6%; Score 22; DB 14; Length 215;  
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYKF 29

Db 90 FTASKNGTYKF 100

## RESULT 9

US-10-775-165-4  
; Sequence 4, Application US/10775165  
; Publication No. US20040138443A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: TI-R Ligand III  
; FILE REFERENCE: PF357C1  
; CURRENT APPLICATION NUMBER: US/10/775,165  
; CURRENT FILING DATE: 2004-02-11

```

; PRIOR APPLICATION NUMBER: US/10/215,088
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/030,847
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/039,483
; PRIOR FILING DATE: 1997-02-28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-165-4

Query Match      38.6%; Score 22; DB 16; Length 215;
Best Local Similarity 36.4%; Pred. No. 8.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXGXYYF 29
Db      89 FTASKNGTYKF 99

RESULT 10
US-10-276-774-2150
; Sequence 2150, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2150
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2150

Query Match      38.6%; Score 22; DB 15; Length 219;
Best Local Similarity 36.4%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXGXYYF 29
Db      93 FTASKNGTYKF 103

RESULT 11
US-10-282-122A-67003
; Sequence 67003, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67003
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-67003

Query Match      38.6%; Score 22; DB 15; Length 296;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXGXYYF 29
Db      276 FTAKVGSYYF 286

RESULT 12
US-10-282-122A-77397
; Sequence 77397, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27/257,931
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77397
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77397

Query Match          38.6%; Score 22; DB 15; Length 302;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXXYP 29
DB      277 FTKAVTGATYFF 287

RESULT 13
US-10-437-963-159816
; Sequence 159816, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159816
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59155C.1.pep
US-10-437-963-159816

Query Match          38.6%; Score 22; DB 16; Length 374;
Best Local Similarity 36.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXXYP 29
DB      207 PSANGDGTYSF 217

RESULT 14
US-10-425-114-68352
; Sequence 68352, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B

US-10-425-115-274242
; Sequence 274242, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 274242
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_181699C.1.pep
US-10-425-115-274242

Query Match          38.6%; Score 22; DB 17; Length 636;
Best Local Similarity 36.4%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXXYP 29
DB      500 FVASSAGLYRF 510

Search completed: December 25, 2004, 09:22:22
Job time : 59.5443 secs
```

This Page Blank (uspto)